Appunti di Algebra

(*versione provvisoria incompleta*)

A. Carboni

Ottobre 2000

Capitolo 1

Funzioni e Contare

1.1 Funzioni e composizione di funzioni

Non definiamo che cosa è un insieme, lasciando ad ognuno immaginare esempi particolari di insiemi: l'insieme dei punti di una retta, l'insieme dei numeri razionali, l'insieme di tutti i movimenti rigidi del piano, l'insieme dei simboli di un linguaggio, e così via. Indicheremo gli insiemi con lettere maiuscole $A, B, \ldots, X, Y$, ecc., mentre scriveremo $a \in A$, $x \in X$, ecc. per indicare che un certo elemento a (risp. x) appartiene all'insieme A (risp. X).

La più importante nozione che si considera per gli insiemi è quella di *funzione* (o anche *applicazione* o *morfismo*): una funzione da un insieme A ad un insieme B è una qualunque legge o regola f che associa ad *ogni* elemento di A *uno ed un solo* elemento di B. Gli insiemi A e B sono chiamati rispettivamente *dominio* e *codominio* della funzione f. Per indicare che f è una funzione di dominio A e di codominio B, scriveremo:

$$f\colon A \longrightarrow B$$

e, se $a \in A$, indicheremo con $f(a)$ l'unico elemento di B che f associa ad a. L'elemento $f(a)$ di B si chiama *immagine di a nella funzione f* o anche *valore assunto da f sull'elemento a di A*. Per le funzioni vale il seguente

Principio di estensionalità: due funzioni $f, g\colon A \longrightarrow B$ sono uguali se e solo se per ogni elemento $a \in A$ si ha $f(a) = g(a)$.

L' operazione fondamentale definita per le funzioni è la *composizione*: se

$$f\colon A \longrightarrow B \quad , \quad g\colon B \longrightarrow C$$

sono due funzioni tali che *il codominio della prima è uguale al dominio della seconda*, allora la regola che consiste nell'associare ad ogni elemento $a \in A$ l'elemento

$$g(f(a))$$

di C definisce una funzione

$$gf\colon A \longrightarrow C$$

che chiamiamo *composizione di f e g* e che potremo leggere anche come "*g segue f*".

Se $h\colon C \longrightarrow D$ è un'altra funzione componibile con g, allora le due composizioni $(hg)f$ e $h(gf)$ esistono ed il principio di estensionalità assicura che

$$(hg)f = h(gf)\,,$$

poichè entrambi i membri valgono $h(g(f(a)))$ per ogni $a \in A$, cioè che vale la *associatività della composizione di funzioni.*

L'altra importante proprietà della composizione di funzioni è *l'esistenza della funzione identità*, cioè per ogni insieme X l'esistenza di una funzione:

$$1_X\colon X \longrightarrow X$$

definita da

$$1_X(x) = x\,,$$

con la proprietà che *per ogni* funzione $f\colon Y \longrightarrow X$ e *per ogni* funzione $g\colon X \longrightarrow Z$ si ha:

$$1_X f = f \qquad \text{e} \qquad g1_X = g\,,$$

come si dimostra facilmente, ancora una volta usando il principio di estensionalità per le funzioni.

Si osservi che la funzione identità è *unica*, nel senso che se esistesse un'altra funzione $1'_X$ con le proprietà di identità, allora $1_X = 1'_X$; infatti: $1_X 1'_X = 1'_X$, perchè 1_X è identità a sinistra e $1_X 1'_X = 1_X$, perchè $1'_X$ è identità a destra.

1.2 Mono, epi, iso, endo e auto (morfismi)

Osserviamo che per dimostrare l'unicità della funzione identità su X sarebbe bastato richiedere le proprietà di identità solo per le funzioni $f, g: X \longrightarrow X$. Tali funzioni, quelle cioè il cui dominio è uguale al codominio, sono dette *endofunzioni* o *endomorfismi*.

Un altro importante concetto, che si esprime solo in termini di composizione di funzioni e delle funzioni identità è quello di *isomorfismo*. Una funzione

$$f: X \longrightarrow Y$$

è un *isomorfismo*, o una *funzione invertibile*, se esiste una funzione

$$g: Y \longrightarrow X$$

tale che

$$gf = 1_X \qquad \text{e} \qquad fg = 1_Y\,,$$

cioè tale che

$$g(f(x)) = x \qquad \text{e} \qquad f(g(y)) = y$$

per ogni $x \in X$ e $y \in Y$. Anche in questo caso si dimostra l'*unicità* di una tale funzione g; se esistesse un'altra funzione $g': Y \longrightarrow X$ per cui $g'f = 1_X$ e $fg' = 1_Y$, allora $g = g'$:

$g = 1_X g = (g'f)g = g'(fg) = g'1_Y = g'$.

Si osservi che abbiamo in realtà dimostrato una proprietà più generale, perchè abbiamo usato solo due delle quattro equazioni nelle ipotesi. Si noti anche che nella dimostrazione si usa la associatività della composizione di funzioni e, come vedremo, in modo essenziale. La funzione g con le proprietà di essere inversa a f, quando esiste, è detta *funzione inversa* di f e viene usualmente denotata con f^{-1}. Infine, se $f: X \longrightarrow Y$ è un isomorfismo, anche f^{-1} è un isomorfismo, essendo f stessa la sua funzione inversa e dunque $(f^{-1})^{-1} = f$. Un endomorfismo di un insieme X che sia anche un isomorfismo viene detto *automorfismo* o anche *permutazione* di X.

Teorema 1.2.1 *Una funzione $f: X \longrightarrow Y$ è un isomorfismo se e solo se vale la proprietà: per ogni $y \in Y$ esiste un solo $x \in X$ tale che $f(x) = y$.*

DIMOSTRAZIONE. Supponiamo che f sia un isomorfismo, cioè che esista una (unica) funzione $f^{-1}: Y \longrightarrow X$ tale che $ff^{-1} = 1_Y$ e $f^{-1}f = 1_X$. Dato un $y \in Y$, definiamo x come l'elemento $x = f^{-1}(y)$; dobbiamo mostrare che $f(x) = y$ e che x è l'unico elemento con tale proprietà. In base alla definizione di x e per il fatto che $ff^{-1} = 1_Y$, si ha: $f(x) = f(f^{-1}(y)) = y$; se x' è un altro elemento per cui $f(x') = y$, allora: $x = f^{-1}(y) = f^{-1}(f(x')) = x'$, poichè $f^{-1}f = 1_X$.

Viceversa, se $f: X \longrightarrow Y$ è una funzione che soddisfa la proprietà espressa dal teorema, definiamo $f^{-1}: Y \longrightarrow X$ come la funzione che ad ogni $y \in Y$ associa l'elemento $f^{-1}(y)$ dato dall'unico $x \in X$ tale che $f(x) = y$. Dobbiamo far vedere che f^{-1} così definita è tale che $ff^{-1} = 1_Y$ e $f^{-1}f = 1_X$. La prima delle due uguaglianze segue direttamente dalla definizione di f^{-1}, mentre la seconda si ottiene dalla condizione di unicità. ■

Questo teorema si può enunciare anche dicendo che un isomorfismo $f: X \longrightarrow Y$ realizza una *corrispondenza biunivoca* tra gli elementi di X e quelli di Y.

Un'ultima importante proprietà degli isomorfismi è la seguente: se $f: A \longrightarrow B$ e $g: B \longrightarrow C$ sono due isomorfismi, allora la composizione

$$A \xrightarrow{\quad f \quad} B \xrightarrow{\quad g \quad} C$$

è un isomorfismo, perchè si dimostra che la sua funzione inversa è

$$C \xrightarrow{\quad g^{-1} \quad} B \xrightarrow{\quad f^{-1} \quad} A\,.$$

Infatti:

$$(gf)(f^{-1}g^{-1}) = ((gf)f^{-1})g^{-1} = (g(ff^{-1}))g^{-1} =$$

$$= (g1_B)g^{-1} = gg^{-1} = 1_C$$

e similmente $(f^{-1}g^{-1})(gf) = 1_A$.

Si noti infine che la funzione identità $1_X: X \longrightarrow X$ è un isomorfismo e dunque un automorfismo, essendo la funzione inversa ancora l'identità.

Analizzando la proprietà espressa dal precedente teorema, si può constatare che essa è in effetti la congiunzione di due distinte proprietà che una funzione $f: X \longrightarrow Y$ può avere:
(E): *per ogni* $y \in Y$ *esiste un* $x \in X$ *tale che* $f(x) = y$*;*
(M): *per ogni coppia di elementi* x *e* x' *di* X*, se* $f(x) = f(x')$*, allora* $x = x'$*.*

Le funzioni che soddisfano solo la proprietà (E) vengono dette *funzioni suriettive* (o anche *epimorfismi*, o più semplicemente *epi*), mentre quelle che soddisfano la proprietà (M) vengono dette *funzioni iniettive* (o anche *monomorfismi*, o più semplicemente *mono*). Il precedente teorema può dunque essere enunciato anche nel modo seguente: *una funzione è un isomorfismo se e solo se è contemporaneamente iniettiva e suriettiva.*

1.3 Esercizi

1. Siano
$$X \xrightarrow{\quad f \quad} Y \xrightarrow{\quad g \quad} Z$$
due funzioni componibili; si dimostri che se f e g sono entrambe iniettive (risp. suriettive), allora la composizione gf è iniettiva (risp. suriettiva).

2. Con le notazioni dell'esercizio precedente, si dimostri che se la composizione gf è iniettiva, allora f è iniettiva e che se gf è suriettiva, allora g è suriettiva.

3. Si dimostri che una funzione $f: X \longrightarrow Y$ è iniettiva se e solo se vale la seguente proprietà: per ogni coppia di funzioni $x, y: U \longrightarrow X$, se $fx = fy$ allora $x = y$.

4. Si provi che f è suriettiva se e solo se per ogni coppia di funzioni $u, v: Y \longrightarrow Z$, se $uf = vf$ allora $u = v$.

1.4 Insiemi finiti

Se n è un numero intero, indichiamo con $[n]$ l'insieme finito standard con n elementi $[n] = \{1, 2, ..., n\}$; diremo che un insieme X è un *insieme finito con n elementi* se esiste una corrispondenza biunivoca

$$x\colon [n] \longrightarrow X$$

tra l'insieme standard con n elementi e X; una tale corrispondenza biunivoca è anche detta *enumerazione* di X e viene spesso indicata semplicemente con $\{x_1, x_2, ..., x_n\}$; in tal caso diremo anche che n è la *cardinalità* di X e scriveremo $|X| = n$. Dunque $|[n]| = n$ e, più in generale, esiste una corrispondenza biunivoca $f\colon X \longrightarrow Y$ se e solo se $|X| = |Y|$ (si veda l'esercizio 1.5.2).

Per meglio comprendere le nozioni di isomorfismo, monomorfismo ed epimorfismo introdotte nel precedente paragrafo, cerchiamo di risolvere il seguente problema di *conteggio*: dati due insiemi finiti X e Y, dare una formula esplicita per la cardinalità dei seguenti insiemi finiti:

1. l'insieme costituito da *tutte* le funzioni $f\colon X \longrightarrow Y$, che indicheremo con la notazione esponenziale Y^X,
2. l'insieme $\mathrm{Mono}(X, Y)$ di tutti i monomorfismi $X \longrightarrow Y$,
3. l'insieme $\mathrm{Iso}(X, Y)$ di tutti gli isomorfismi $X \longrightarrow Y$,
4. l'insieme $\mathrm{Epi}(X, Y)$ di tutti gli epimorfismi $X \longrightarrow Y$,

in termini della cardinalità degli insiemi X e Y.

Per quanto riguarda il problema 1, un semplice ragionamento porta a concludere che

$$\boxed{|Y^X| = |Y|^{|X|}}.$$

Infatti, una funzione $f\colon X \longrightarrow Y$ è completamente determinata quando per ogni $x \in X$ sia dato l'elemento $f(x) \in Y$; poichè per ogni $x \in X$ l'elemento $f(x)$ di Y può essere scelto in $|Y|$ modi distinti, si ha la formula precedente. Dunque in particolare $|[n]^{[m]}| = n^m$.

Osserviamo che per i problemi di conteggio è sufficiente assumere che gli insiemi finiti in questione siano gli insiemi finiti standard $[n]$. Cerchiamo dunque di determinare il numero $n_{(m)} = |\mathrm{Mono}(m,n)|$ delle funzioni *iniettive* $f\colon [m] \longrightarrow [n]$. Ora, il valore $f(1)$ di una tale funzione f su 1 può essere scelto in n modi distinti, mentre il valore $f(2)$ su 2 può essere scelto solo in $(n-1)$ modi, *perchè la condizione di iniettività per f implica che il valore già scelto per $f(1)$ non possa più essere assunto da f*. In definitiva, le scelte possibili per i valori di f su 1 e 2 sono $n(n-1)$; così continuando si trova che:

$$\boxed{n_{(m)} = n \cdot (n-1) \cdot (n-2) \cdot \ \cdots \ \cdot (n-m+1)\,.}$$

Si osservi che in particolare si ha $n_{(m)} = 0$, se m è maggiore di n e che in particolare, se $m = n$, si ha:

$$\boxed{n_{(n)} = n! = n \cdot (n-1) \cdot (n-2) \cdot \ \cdots \ \cdot 2 \cdot 1}\,.$$

Dunque il numero $n!$ (che si legge "fattoriale di n" od anche "n fattoriale") conta il numero di tutte le funzioni iniettive $\sigma\colon [n] \longrightarrow [n]$. Ma:

Teorema 1.4.1 *Se X è un insieme finito, allora un endomorfismo $\sigma\colon X \longrightarrow X$ di X è iniettivo se e solo se è suriettivo.*

DIMOSTRAZIONE. Sia σ iniettivo e sia $y \in X$; se $y = \sigma(y)$ allora esiste un x tale che $\sigma(x) = y$ (basta prendere $x = y$); se $y \neq \sigma(y) = x_1$, allora anche $\sigma(y) \neq \sigma(x_1) = x_2$, perchè σ è iniettiva; similmente anche $x_3 = \sigma(x_2)$ è diverso da tutti precedenti; dunque continuando ad applicare σ si ottengono tutti elementi distinti di X; ma poichè X è finito, ad un certo punto si dovrà riottenere anche y come $\sigma(x_k)$; dunque σ è suriettiva.

Viceversa, supponiamo che σ non sia iniettiva; dunque esistono $x \in X$ e $x' \in X$ con $x \neq x'$ e $\sigma(x) = \sigma(x')$. Perciò il numero degli elementi di X della forma $y = \sigma(x)$ deve essere strettamente minore del numero degli elementi di X; dunque σ non può essere suriettiva. ■

Dunque, il numero $n!$ conta il numero degli endomorfismi iniettivi di $[n]$, perciò degli *automorfismi* di n; tale insieme è spesso indicato con S_n ed i suoi elementi sono anche chiamati *permutazioni* di $[n]$; riassumendo si ha:

$$\boxed{|S_n| = n!}.$$

Si osservi che se X è un insieme di cardinalità n, allora il numero delle *enumerazioni* $[n] \longrightarrow X$ di X è ancora $n!$.

Infine, osserviamo che la proprietà degli insiemi finiti espressa nel precedente teorema non contiene alcun riferimento ai numeri naturali, ma solo ad una proprietà delle loro endofunzioni. Essa potrebbe essere presa come *definizione* di insieme finito senza che si debba assumere l'esistenza dei numeri naturali. Il lettore interessato può domandarsi se in presenza dei numeri naturali le due definizioni sono equivalenti e se, non assumendo l'esistenza dei numeri naturali, quante delle proprietà degli insiemi finiti che useremo nel seguito possono essere dimostrate solo a partire dalla definizione espressa nell'enunciato del teorema.

Il problema di contare il numero $\mathrm{Epi}(X, Y)$ degli epimorfismi tra due insiemi finiti X e Y è un pò più complicato e lo affronteremo nei prossimi paragrafi.

1.5 Esercizi

1. Si elenchino esplicitamente gli elementi di Mono([2],[4]), S_2 e S_3.

2. Siano X e Y insiemi finiti. Si provi che se $|X|$ è maggiore di $|Y|$, allora non esistono funzioni iniettive $f\colon X \longrightarrow Y$. Similmente, se $|X|$ è minore di $|Y|$, allora non esistono funzioni suriettive $f\colon X \longrightarrow Y$. In particolare, se $f\colon [n] \longrightarrow [m]$ è un isomorfismo, allora $n = m$.

1.6 Il reticolo dei sottoinsiemi di un insieme

Se X è un insieme, che cosa sia un sottoinsieme di X è del tutto intuitivo: diremo che un insieme U è un *sottoinsieme* di X (o anche che è una *parte* di X) e scriveremo

$$U \subseteq X$$

(che leggeremo "U è contenuto in X") se ogni elemento di U è anche un elemento di X. Naturalmente, se per un altro insieme V si ha $V \subseteq U$, allora si ha anche $V \subseteq X$. Inoltre, per i sottoinsiemi di un insieme X vale il seguente:

Principio di estensionalità: siano $U \subseteq X$ e $V \subseteq X$; allora $U = V$ se e solo se U e V hanno gli stessi elementi, cioè se e solo se per ogni $x \in X$ si ha: $x \in U$ se e solo se $x \in V$.

Indicheremo l'insieme dei sottoinsiemi di X con la notazione

$$\mathbf{P}X.$$

Per meglio comprendere la nozione di sottoinsieme di un insieme, risolviamo il problema di contare il numero dei sottoinsiemi di un insieme finito X. Allo scopo cerchiamo di ridurre il problema ad un problema di conteggio di funzioni del tipo che abbiamo già risolto, mediante la seguente osservazione. Sia $\mathbf{2}$ l'insieme

$$\mathbf{2} = \{0, 1\}$$

dei valori di verità 1 = "vero" e 0 = "falso". Se $U \subseteq X$, possiamo considerare la sua *funzione caratteristica*

$$c_U \colon X \longrightarrow \mathbf{2}$$

definita da:

$$c_U(x) = \begin{cases} 1 & \text{se } x \in U \\ 0 & \text{se } x \notin U \end{cases}$$

($x \notin U$ significa che x non appartiene ad U). Viceversa, ogni funzione $c \colon X \longrightarrow \mathbf{2}$ definisce un unico sottoinsieme $U_c \subseteq X$ tale che $c_{U_c} = c$,

mediante $U_c = \{x \in X | c(x) = 1\}$. Dunque la costruzione della funzione caratteristica di un sottoinsieme stabilisce una *corrispondenza biunivoca* tra l'insieme $\mathbf{P}X$ dei sottoinsiemi di X e l'insieme $\mathbf{2}^X$ delle funzioni $c\colon X \longrightarrow \mathbf{2}$. Perciò

$$|\mathbf{P}X| = |\mathbf{2}^X| = 2^{|X|}.$$

ed in particolare $|\mathbf{P}[n]| = 2^n$. Spesso, come abbiamo fatto per U_c, indicheremo un sottoinsieme di un insieme X definito da una proprietà P degli elementi di X, con la notazione

$$U_P = \{x \in X | P(x)\}$$

da leggersi "l'insieme U degli elementi $x \in X$ per cui la proprietà P è verificata".

Per meglio comprendere la differenza tra la nozione di sottoinsieme e quella di funzione iniettiva, risolviamo il seguente problema di conteggio: *dare una formula esplicita per il numero dei sottoinsiemi di* $[n]$ *aventi* k *elementi.* Per determinare tale numero ragioniamo come segue: se $f\colon [k] \longrightarrow [n]$ è una funzione iniettiva, l'insieme $I(f) = \{f(1), f(2), \ldots, f(k)\}$ dei suoi valori è certamente un sottoinsieme di $[n]$ avente k elementi, poichè f è iniettiva; tuttavia, se $\sigma\colon [k] \longrightarrow [k]$ è una permutazione di $[k]$, la funzione iniettiva $f\sigma\colon [k] \longrightarrow [n]$ che si ottiene per composizione definisce lo stesso sottoinsieme di $[n]$: infatti in $I(f\sigma) = \{f(\sigma(1)), f(\sigma(2)), \ldots, f(\sigma(k))\}$ compaiono ancora tutti e soli gli elementi di $I(f)$; dunque esistono $k!$ funzioni iniettive $[k] \longrightarrow [n]$ che determinano lo *stesso* sottoinsieme di $[n]$ con k elementi. Perciò, dato che ogni sottoinsieme di $[n]$ con k elementi determina una funzione iniettiva $[k] \longrightarrow [n]$, indicando con $\binom{n}{k}$ il numero dei sottoinsiemi di $[n]$ con k elementi e ricordando che abbiamo indicato con $n_{(k)}$ il numero delle funzioni iniettive $[k] \longrightarrow [n]$, si ha che

$$\boxed{k!\binom{n}{k} = n_{(k)}}.$$

Ancora, ricordando la formula per $n_{(k)}$ si ha:

$$\boxed{\binom{n}{k} = \frac{n\cdot(n-1)\cdot\cdots\cdot(n-k+1)}{k!} = \frac{n!}{k!(n-k)!}}.$$

Si osservi che

$$\binom{n}{0} = 1 \quad \text{e} \quad \binom{n}{n} = 1,$$

poichè per ogni insieme X esiste un unico sottoinsieme senza elementi ed un unico sottoinsieme che ha gli stessi elementi di X.

I numeri $\binom{n}{k}$ sono chiamati *coefficienti binomiali*; non è difficile convincersi che essi hanno un carattere combinatorio: essi esprimono il numero di modi possibili di effettuare una scelta di k oggetti distinti su n oggetti dati, in modo che due scelte differiscano per almeno un oggetto.

Sull'insieme $\mathbf{P}X$ delle parti di un insieme X sono definite le usuali operazioni di *intersezione*, *unione* e *complemento*. Se $U \subseteq X$ e $V \subseteq X$, si definiscono i sottoinsiemi:

$$U \cap V = \{x \in X | x \in U \text{ e } x \in V\} \qquad \text{(intersezione)}$$

$$U \cup V = \{x \in X | x \in U \text{ o } x \in V\} \qquad \text{(unione)}$$

$$U^c = X - U = \{x \in X | x \notin U\}. \qquad \text{(complemento)}$$

Due sottoinsiemi che ogni insieme X possiede sempre (*sottoinsiemi impropri*) sono il *sottoinsieme vuoto* $\emptyset$ ed il *sottoinsieme totale*, cioè X stesso.

Usando il principio di estensionalità per i sottoinsiemi può essere un utile esercizio dimostrare le seguenti identità per le operazioni di intersezione, unione e complemento di sottoinsiemi U, V e W di X:

$$(U\cup V)\cup W = U\cup(V\cup W),\ (U\cap V)\cap W = U\cap(V\cap W) \quad \text{(associatività)}$$

$$U \cup V = V \cup U,\ U \cap V = V \cap U \quad \text{(commutatività)}$$

$$U \cup U = U,\ U \cap U = U \quad \text{(idempotenza)}$$

$$\emptyset \cup U = U,\ X \cap U = U \quad \text{(elemento neutro)}$$

$$U \cap (V \cup W) = (U \cap V) \cup (U \cap W) \quad \text{(distributività)}$$

$$U \cup U^c = X,\ U \cap U^c = \emptyset. \quad \text{(complemento)}$$

1.7 Esercizi

1. Sia X un insieme finito e siano U e V due suoi sottoinsiemi; si provi che:
$$|U \cup V| + |U \cap V| = |U| + |V|$$
$$|U^c| = |X| - |U|.$$

2. Si dimostrino le identità:
$$\sum_{i=0}^{n} \binom{n}{i} = 2^n$$
$$\binom{n}{k} = \binom{n}{n-k}$$
$$\binom{n}{k}\binom{k}{m} = \binom{n-m}{k-m}\binom{n}{m},$$
con argomenti di carattere puramente combinatorio usando l'interpretazione dei coefficienti binomiali come numero dei sottoinsiemi con assegnata cardinalità (si usa la tecnica di mostrare che entrambi i membri dell'uguaglianza contano gli stessi sottoinsiemi).

3. Un torneo a k turni è una gara in cui ogni squadra incontra ogni altra esattamente k volte. Si determini il numero $P(n,k)$ delle partite da giocare tra n squadre in un torneo a k turni.

1.8 La formula del binomio di Newton

Dati due insiemi X e Y si definisce la loro *somma* (*disgiunta*)
$$X + Y$$
come l'insieme i cui elementi sono quelli di X e quelli di Y, con l'ulteriore clausola che se un elemento appartiene ad entrambi gli insiemi *esso va contato due volte*, una come appartenente ad X ed una come

appartenente ad Y. Dunque, in particolare, se X e Y sono insiemi finiti si ha:

$$|X+Y| = |X| + |Y|.$$

Se gli insiemi finiti X e Y sono gli insiemi finiti standard $[n]$ e $[m]$, allora $|[n]+[m]| = |[n+m]| = n+m$. La nozione di somma $X+Y$ di due insiemi X e Y può essere è data in modo formalmente preciso nel modo seguente: esistono due funzioni

$$i_X : X \longrightarrow X+Y \quad , \quad i_Y : Y \longrightarrow X+Y$$

(dette *iniezioni*), che soddisfano la seguente proprietà (universale):
(P): *per ogni insieme Z e per ogni coppia di funzioni*

$$f : X \longrightarrow Z \quad , \quad g : Y \longrightarrow Z$$

esiste una ed una sola funzione

$$\binom{f}{g} : X+Y \longrightarrow Z$$

tale che $\binom{f}{g} i_X = f$ *e* $\binom{f}{g} i_Y = g$.

In altre parole, la composizione con le iniezioni definisce una *corrispondenza biunivoca* tra le funzioni

$$X+Y \longrightarrow Z$$

e le *coppie ordinate* di funzioni

$$\langle X \longrightarrow Z, Y \longrightarrow Z \rangle.$$

Dunque, in particolare, se X, Y e Z sono insiemi finiti si ha:

$$|Z^{X+Y}| = |Z^X| \cdot |Z^Y|.$$

Descrivere le funzioni che hanno per dominio una somma disgiunta è semplice: segue direttamente dalla definizione di somma. Un po' più complicato è invece descrivere le funzioni che hanno per *codominio* una

somma. Se $V \subseteq X$ e $W \subseteq X$ sono due sottoinsiemi di X che siano *disgiunti* (cioè $V \cap W = \emptyset$), allora

$$V \cup W = V + W.$$

Dunque, se X è un insieme finito e $U \subset X$ è un suo sottoinsieme, allora per ogni insieme finito Z si ha:

$$|Z^X| = |Z^U| \cdot |Z^{U^c}|.$$

Se $f: T \longrightarrow X + Y$ è una funzione da un insieme T ad una somma di insiemi $X + Y$, allora detti U e V i sottoinsiemi di T definiti da

$$U = \{t \in T | f(t) \in X\} \quad , \quad V = \{t \in T | f(t) \in Y\}$$

si ha che V è il complementare U^c di U e che f definisce una coppia di funzioni $f_X: U \longrightarrow X$ e $f_Y: U^c \longrightarrow Y$. Viceversa, dato un sottoinsieme U di T, una coppia di funzioni $f_X: U \longrightarrow X$ e $f_Y: U^c \longrightarrow Y$ definisce un'unica funzione $f: T \longrightarrow X + Y$. Dunque assegnare una funzione $f: T \longrightarrow X + Y$ *equivale* ad assegnare un sottoinsieme $U \subseteq T$ ed una coppia di funzioni $U \longrightarrow X$ e $U^c \longrightarrow Y$. Ne segue che se $[n]$, $[a]$ e $[b]$ sono tre insiemi finiti standard, allora le funzioni $[n] \longrightarrow [a] + [b]$ possono essere contate nel modo seguente: si contano dapprima le funzioni $[n] \longrightarrow [a]+[b]$ che mandano un sottoinsieme assegnato $U \subseteq [n]$ di cardinalità k in $[a]$, il cui numero in base al ragionamento precedente risulta essere

$$a^k b^{n-k};$$

poi si osserva che, dato che il numero dei sottoinsiemi di cardinalità k è $\binom{n}{k}$, il numero delle funzioni che mandano un *arbitrario* sottoinsieme di $[n]$ di cardinalità k in $[a]$ è

$$\binom{n}{k} a^k b^{n-k};$$

è chiaro che per ottenere il numero di *tutte* le funzioni $[n] \longrightarrow [a] + [b]$ bisogna sommare tali numeri rispetto a tutte le possibili cardinalità dei

sottoinsiemi di n; dunque

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

Questa formula è chiamata *formula del binomio di Newton (per i numeri naturali)*. Vedremo in seguito come essa si estende a tutti i numeri interi a e b e, più in generale, ad ogni coppia di elementi permutabili a e b di ogni anello.

1.9 Esercizi

1. Si definisca esplicitamente un isomorfismo $[n] + [m] \longrightarrow [n+m]$.

2. Si dimostri la seguente identità:

$$\binom{n+1}{k+1} = \binom{n}{k+1} + \binom{n}{k};$$

(suggerimento: si ripartiscano i sottoinsiemi $U \subseteq [n+1]$ in due classi determinate dalla condizioni $(n+1) \in U$ oppure $(n+1) \notin U$ e si contino i sottoinsiemi di ciascuna classe).

3. Si provi che se X e Y sono due insiemi finiti, allora $|\mathbf{P}(X+Y)| = |\mathbf{P}X| \cdot |\mathbf{P}Y|$.

4. Si provi la seguente identità ("convoluzione di Vandermonde"):

$$\binom{m+n}{k} = \sum_{i=0}^{k} \binom{m}{i} \binom{n}{k-i}.$$

(Suggerimento: ogni sottoinsieme di $[m] + [n]$ con k elementi si può ottenere scegliendo un sottoinsieme di $[m]$ con i elementi ed un sottoinsieme di $[n]$ con $(k-i)$ elementi).

5. Si provi la seguente identià:

$$\binom{n+1}{k+1} = \sum_{i=k}^{n} \binom{i}{k}.$$

6. Si provi che

$$\sum_{i=0}^{k}(-1)^i\binom{k}{i} = \delta_{k,0}\,,$$

essendo $\delta_{k,h}$ il "simbolo di Kronecker": $\delta_{k,h} = 0$ se $k \neq h$ e $\delta_{k,h} = 1$ se $k = h$. (Suggerimento: assumendo che, come è stato preannunciato, la formula del binomio di Newton è valida anche per gli interi sia positivi che negativi, la si applichi a $(1-1)^k$ e si osservi che $(1-1)^k = 0$ se $k \neq 0$, mentre $(1-1)^0 = 1$).

1.10 Funzioni monotòne

Gli insiemi standard $[n]$ hanno un ordine naturale sui loro elementi:

$$1 < 2 < 3 < \cdots < n$$

e accade in pratica di dover considerare funzioni $f\colon [n] \longrightarrow [m]$ che rispettano tale ordine (*funzioni monotòne* (*non decrescenti*)), cioè funzioni che soddisfano:

$$\text{se } \; i < j \quad \text{allora} \quad f(i) \leq f(j)$$

per ogni $i, j \in [n]$ (il simbolo "$\leq$ " significa "minore o uguale "). Ancora, poniamoci il problema di contare il numero $D_{n,m}$ delle funzioni monotòne non decrescenti $f\colon [n] \longrightarrow [m]$.

Lemma 1.10.1

$$D_{n,m} = \binom{m+n-1}{n}\,.$$

DIMOSTRAZIONE. Se $f\colon [n] \longrightarrow [m]$ è una funzione monotòna, consideriamo la funzione $g\colon [n] \longrightarrow [m+n-1]$ così definita:

$$g(1) = f(1), \quad g(2) = f(2)+1, \quad \ldots \quad , g(n) = f(n)+n-1.$$

Tale funzione g è ancora monotòna, ma in più è anche *iniettiva*. Viceversa, data una funzione $g\colon [n] \longrightarrow [m+n-1]$ che sia monotòna ed

iniettiva (dunque $g(1) < g(2) < \cdots < g(n)$), si ottiene una funzione monotòna $f\colon [n] \longrightarrow [m]$ mediante:

$$f(1) = g(1), \quad f(2) = g(2) - 1, \quad \ldots \quad , f(n) = g(n) - (n-1).$$

Poichè è chiaro che le due corrispondenze sono inverse l'una all'altra, si ha che il numero delle funzioni monotòne $f\colon [n] \longrightarrow [m]$ è uguale al numero delle funzioni monotòne e *iniettive* $g\colon [n] \longrightarrow [m+n-1]$. Contiamo dunque in generale quante sono le funzioni monotòne ed iniettive $[p] \longrightarrow [q]$. Ogni funzione monotòna ed iniettiva $h\colon [p] \longrightarrow [q]$ determina un sottoinsieme di $[q]$ con p elementi $\{h(1), h(2), \ldots, h(p)\}$, i cui elementi hanno lo stesso ordine che hanno in $[q]$ perchè h è monotòna. Viceversa, dato un sottoinsieme di $[q]$ con p elementi, ordinando i suoi elementi con l'ordine di $[q]$ si ottiene una funzione monotòna ed iniettiva $[p] \longrightarrow [q]$; dunque le funzioni monotòne ed iniettive $[p] \longrightarrow [q]$ sono tante quante i sottoinsiemi di $[q]$ con p elementi, cioè $\binom{q}{p}$. ■

I numeri $D_{n,m}$ hanno un significato combinatorio: essi contano le *combinazioni con ripetizione* di n oggetti su m oggetti dati. Ricordiamo che una combinazione con ripetizione di n oggetti su m oggetti dati è una famiglia $D = \{d_i\}_{i\in[n]}$ di n elementi $d_i \in [m]$, non necessariamente distinti; due tali famiglie definiscono la stessa disposizione con ripetizione se differiscono solo per l'ordine. In altre parole, una combinazione con ripetizione D è semplicemente una funzione $d\colon [n] \longrightarrow [m]$, pur di convenire che due funzioni $d, d'\colon [n] \longrightarrow [m]$ definiscono la stessa combinazione se esiste una permutazione $\sigma\colon [n] \longrightarrow [n]$ tale che $d\sigma = d'$. Per mostrare che i numeri $D_{n,m}$ contano le combinazioni con ripetizione di n oggetti su m oggetti dati, ragioniamo nel modo seguente: una funzione monotòna $d\colon [n] \longrightarrow [m]$ certamente definisce una combinazione $D = \{d_i = d(i)\}_{i\in[n]}$; d'altra parte, se $D = \{d_i\}_{i\in[n]}$ è una combinazione, si può sempre trovare una permutazione σ di $[n]$ tale che la famiglia $D' = \{d_{\sigma(i)}\}_{i\in[n]}$ (che definisce la *stessa* combinazione) sia *ordinata*, cioè $d_{\sigma(i)} \leq d_{\sigma(j)}$ se $i < j$ e quindi che definisca una funzione monotòna $d'\colon [n] \longrightarrow [m]$.

1.11 Esercizi

1. Si elenchino esplicitamente gli elementi di $D_{3,2}$, $D_{3,3}$.

2. Diremo che un sottoinsieme $I \subseteq [n]$ è un *intervallo* di $[n]$ se I è costituito da elementi *consecutivi* di $[n]$. Si provi che il numero delle partizioni di $[n]$ in k intervalli è

$$\binom{n-1}{k-1.}$$

(Suggerimento: una partizione di $[n]$ in k intervalli è univocamente determinata dalla scelta di un sottoinsieme di $[n]$ con $(k-1)$ elementi che non contenga n).

3. Si provi che se $f\colon [n] \longrightarrow [m]$ è una funzione monotòna, allora per ogni $j \in Im(f)$ i sottoinsiemi $I_j = f^*(j) = \{i \in [n] | f(i) = j\}$ costituiscono una partizione di $[n]$ in k intervalli, essendo $k = |Im(f)|$.

4. Si dimostri la seguente identità in modo combinatorio, usando i precedenti esercizi:

$$D_{n,m} = \sum_{k=1}^{n} \binom{n-1}{k-1} \binom{m}{k}.$$

(Si osservi che, conoscendo la formula per $D_{n,m}$, tale identità è un caso particolare della formula di convoluzione di Vandermonde (esercizio 1.9.4).

5. Denotando con $x^{(n)}$ la cardinalità dell'insieme Mono$([n][x+n-1])$, si dimostri l'identità:

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

6. Si provi che il numero dei modo di ripartire n oggetti distinguibili (ad esempio numerati) in m pile ordinate non vuote è

$$n! \binom{n-1}{m-1}.$$

7. Si provi che la matrice quadrata

$$L = \left[\frac{i!}{k!} \binom{i-1}{k-1} \right]$$

è invertibile e che la sua inversa ha per elementi gli interi

$$l(i,k) = (-1)^{i+k} \frac{i!}{k!} \binom{i-1}{k-1}.$$

Si deduca che:

$$m_{(n)} = \sum_{k=1}^{n} l(n,k) m^{(k)}.$$

Gli interi $l(i,k)$ sono chiamati numeri di Lah.

1.12 Prodotti

La scelta di usare la stessa notazione (lettere minuscole) per indicare sia le funzioni sia gli elementi è giustificata dall'esistenza di un *insieme con un solo elemento*, che indicheremo con il simbolo "$*$". È chiaro infatti che una funzione $x: * \longrightarrow X$ di dominio $*$ altro non è che un elemento $x \in X$. È anche chiaro che l'insieme $*$ soddisfa la seguente proprietà:
(!): *per ogni insieme X esiste una sola funzione* $!_X: X \longrightarrow *$.

Proprietà del tipo della proprietà (!) per un particolare insieme o per un particolare "diagramma" sono chiamate "proprietà universali". Esse fanno riferimento a tutti gli altri insiemi e funzioni e determinano l'insieme o il diagramma in questione a meno di un unico isomorfismo che soddisfa opportune condizioni. Vedremo in seguito una loro precisa definizione.

Un altro esempio di proprietà universale è quella soddisfatta dal *prodotto cartesiano di insiemi*. Dati due insiemi A e B, l'insieme

$$A \times B$$

i cui elementi sono le *coppie ordinate* $\langle a, b \rangle$ di elementi, il primo di A ed il secondo di B, è detto *prodotto cartesiano* degli insiemi A e B (in questo ordine!). Esistono due funzioni:

$$p_A: A \times B \longrightarrow A \quad , \quad p_B: A \times B \longrightarrow B$$

(dette *proiezioni*), definite da

$$p_A\langle a, b\rangle = a \quad \text{e} \quad p_B\langle a, b\rangle = b$$

che soddisfano la seguente proprietà (universale):
(P): *per ogni insieme* X *e per ogni coppia di funzioni*

$$f\colon X \longrightarrow A \quad , \quad g\colon X \longrightarrow B$$

esiste una ed una sola funzione

$$\langle f, g\rangle : X \longrightarrow A \times B$$

tale che $p_A\langle f, g\rangle = f$ *e* $p_B\langle f, g\rangle = g$.

In altre parole, la composizione con le proiezioni definisce una *corrispondenza biunivoca* tra le funzioni

$$X \longrightarrow A \times B$$

e le *coppie ordinate* di funzioni

$$\langle X \longrightarrow A, X \longrightarrow B\rangle.$$

Per dimostrare tale proprietà basta definire la funzione $\langle f, g\rangle$ come

$$\langle f, g\rangle(x) = \langle f(x), g(x)\rangle.$$

Sarà utile nel seguito considerare le seguenti funzioni, la cui esistenza e unicità, nonchè il fatto che sono *isomorfismi*, si potrebbe dimostrare solo facendo uso della proprietà (P):

1. L' isomorfismo $\langle 1_A, !_A\rangle\colon A \longrightarrow A \times *$;

2. Se A, B e C sono tre insiemi, si possono considerare i prodotti cartesiani $(A \times B) \times C$ e $A \times (B \times C)$; la funzione

   $$\alpha\colon (A \times B) \times C \longrightarrow A \times (B \times C)$$

   definita da $\alpha\langle\langle a, b\rangle, c\rangle = \langle a, \langle b, c\rangle\rangle$ è un isomorfismo, detto "isomorfismo di associatività del prodotto cartesiano". Spesso ignoreremo questo isomorfismo e scriveremo semplicemente

   $$A \times B \times C$$

per indicare l'insieme i cui elementi sono le *terne ordinate*

$$\langle a, b, c\rangle$$

di elementi, il primo di A, il secondo di B ed il terzo di C. È chiaro come tutto ciò si generalizza al prodotto cartesiano di un numero finito di insiemi.

Nel caso di $A = B$, indicheremo con A^2 l'insieme $A \times A$ delle coppie ordinate di elementi di A. Più in generale, indicheremo con A^n (dove n indica un numero intero) l'insieme $A \times A \times \cdots \times A$ (n volte) delle n-uple ordinate di elementi di A. Si osservi che A^n altro non è che l'insieme $A^{[n]}$ delle funzioni $[n] \longrightarrow A$ e che dunque $A^0 = *$, poichè, per ogni insieme A, esiste un'unica funzione

$$\emptyset \longrightarrow A\,.$$

3. Per ogni coppia di insiemi A e B si ha l'isomorfismo

$$\sigma\colon A \times B \longrightarrow B \times A$$

definito da $\sigma(a, b) = \langle b, a\rangle$, detto "isomorfismo di simmetria".

Il prodotto cartesiano di insiemi si estende alle funzioni nel modo seguente: se $f\colon A \longrightarrow B$ e $g\colon C \longrightarrow D$ sono due funzioni, allora si definisce la funzione:

$$f \times g\colon A \times C \longrightarrow B \times D$$

come $(f \times g)(a, c) = \langle f(a), g(c)\rangle$. È facile vedere che si hanno le due seguenti proprietà:

$$1_A \times 1_B = 1_{A\times B}$$

e, se $h\colon B \longrightarrow E$, $k\colon D \longrightarrow H$ sono altre due funzioni, la prima componibile con f e la seconda componibile con g, allora (esercizio):

$$(h \times k)(f \times g) = (hf) \times (kg).$$

L'ultima importante proprietà del prodotto cartesiano è la sua relazione con l'insieme delle funzioni. Ricordiamo che per ogni coppia di insiemi X e B abbiamo indicato con

$$B^X$$

l'insieme di tutte le funzioni $X \longrightarrow B$. La relazione fondamentale che lega prodotto cartesiano ed insieme delle funzioni è la "*chiusura cartesiana*" (anche chiamata "λ-*conversione*"): per ogni insieme A, data una funzione in "due variabili"

$$f\colon A \times X \longrightarrow B$$

si ottiene una funzione in una variabile ("*aggiunta esponenziale*"):

$$\lambda f\colon A \longrightarrow B^X$$

mediante

$$\lambda f(a)(x) = f(a, x);$$

cioè $\lambda f(a)$ è la funzione $X \longrightarrow B$ che si ottiene da f tenendo fisso l'elemento $a \in A$. Viceversa, se $g\colon A \longrightarrow B^X$ è una funzione di codominio B^X, allora possiamo associare a g una funzione

$$g^t\colon A \times X \longrightarrow B$$

(detta "*trasposta di* g"), definendo $g^t(a, x) = g(a)(x)$, cioè associando ad ogni coppia $\langle a, x\rangle$ il valore che la funzione $g(a)\colon X \longrightarrow B$ assume sull'elemento $x \in X$. È immediato constatare che λ è un isomorfismo tra l'insieme delle funzioni

$$A \times X \longrightarrow B$$

e l'insieme delle funzioni

$$A \longrightarrow B^X\,.$$

In particolare, la trasposta della funzione identità $B^X \longrightarrow B^X$, definisce una funzione

$$val_{B,X}\colon B^X \times X \longrightarrow B$$

detta "valutazione", poichè in accordo con la definizione di trasposta, è definita da $val_{B,X}(f, x) = f(x)$.

1.13 Esercizi

1. Si provi che se X e Y sono insiemi finiti, allora

$$\boxed{|X \times Y| = |X| \cdot |Y|}\,.$$

In particolare, $|[n] \times [m]| = nm$.

2. Se X, Y e Z son tre insiemi, si provi che la funzione

$$\begin{pmatrix} i_X \times 1_Z \\ i_Y \times 1_Z \end{pmatrix} : X \times Z + Y \times Z \longrightarrow (X + Y) \times Z$$

è un isomorfismo (*proprietà distributiva del prodotto rispetto alla somma*).

3. Se X è un insieme, si consideri il sottoinsieme del prodotto $X \times X$ definito da

$$\Delta_X = \{\langle x, y\rangle \mid x = y\} \subseteq X \times X\,.$$

Tale sottoinsieme è detto *diagonale* di X. Si provi che la sua funzione caratteristica $\delta_X : X \times X \longrightarrow \mathbf{2}$ è il "simbolo di Kronecker" (si veda l'esercizio 6 del paragrafo 1.9).

4. Si provi che il numero delle m-uple *ordinate* $\langle x_1, x_2, \ldots, x_m\rangle$ di interi non negativi x_i tali che

$$\sum_{i=1}^{m} x_i = n$$

(con n intero non negativo) è $D_{n,m}$. (Suggerimento: data una funzione $x : [n] \longrightarrow [m]$ si ottiene una soluzione ponendo

$$x_i = |x^*(i)| = |\{j \in [m] | x(j) = i\}|;$$

inoltre, una permutazione σ di n definisce la *stessa* soluzione, perchè per ogni $i \in [m]$ si ha che σ induce una corrispondenza biunivoca $(x\sigma)^*(i) \simeq x^*(i)$, dunque $|(x\sigma)^*(i)| = |x^*(i)|$; con lo stesso argomento del lemma sul numero delle funzioni monotòne, si vede che le soluzioni sono in corrispondenza biunivoca con le funzioni monotòne $[n] \longrightarrow [m]$, perciò il loro numero è $D_{n,m}$). Si osservi che ciò prova anche che $D_{n,m}$ è il numero dei monomi di grado n nell'anello dei polinomi in m variabili.

5. Si determini il numero delle m-uple ordinate $\langle x_1, x_2, \ldots, x_m \rangle$ di interi *positivi* x_i tali che

$$\sum_{i=1}^{m} x_i = n,$$

essendo n un intero positivo.

6. Si descriva esplicitamente un isomorfismo $f\colon [nm] \longrightarrow [n] \times [m]$. (Suggerimento: si definisca un ordinamento in $[n] \times [m]$ partendo dalla coppia $\langle 1, 1 \rangle$ e definendo induttivamente la coppia successiva $s(i, j)$ secondo la regola:

$$s(i,j) = \begin{cases} \langle i, j+1 \rangle & \text{se } j < m \\ \langle i+1, 1 \rangle & \text{se } j = m \end{cases}$$

ad esempio: $\langle 1,1 \rangle, \langle 1,2 \rangle, \langle 1,3 \rangle, \cdots, \langle 2,1 \rangle, \langle 2,2 \rangle, \langle 2,3 \rangle, \cdots$; tale ordinamento è detto *lessicografico*. La funzione f è allora definita da $f(i) = s^{i-1}(1,1)$, dove $s^0(1,1) = \langle 1,1 \rangle$ e $s^k(1,1) = s[s^{k-1}(1,1)]$).

7. Si definisca un ordinamento sulle coppie di naturali, partendo dalla coppia $\langle 0, 0 \rangle$ e definendo induttivamente la coppia successiva $s(i, j)$ secondo la regola:

$$s(i,j) = \begin{cases} \langle 0, i+1 \rangle & \text{se } j = 0 \\ \langle i+1, j-1 \rangle & \text{se } j \neq 0 \end{cases}$$

si provi che l'ordinamento così definito induce una *corrispondenza biunivoca* $f\colon \mathbf{N} \longrightarrow \mathbf{N} \times \mathbf{N}$, mediante $f(n) = s^n(0,0)$, con $s^0(0,0) = \langle 0,0 \rangle$ e $s^n(0,0) = s[s^{n-1}(0,0)]$. Tale ordinamento è detto "*procedimento diagonale di Cantor*".

8. Dati n insiemi $X_1, X_2, \ldots, X_n$ si conti il numero di prodotti ottenibili dagli n insiemi dati, *nell'ordine dato*, che siano isomorfi tra loro solo in virtù degli isomorfismi di associatività, dei loro prodotti con l'identità e delle composizioni di tali isomorfismi. Così, ad esempio, per $n = 3$ il loro numero è evidentemente 2, mentre per $n = 4$ è 5: omettendo per semplicità il simbolo "$\times$", si hanno i cinque prodotti $(((X_1X_2)X_3)X_4)$, $((X_1(X_2X_3))X_4)$,

$(X_1((X_2X_3)X_4))$, $(X_1X_2)(X_3X_4)$, $(X_1(X_2(X_3X_4)))$. Indicando tale numero con a_n (il risultato non dipende dagli insiemi scelti, ma solo dal loro numero), non è difficile provare la seguente formula induttiva per i numeri a_n (che sono detti *numeri di Catalano*), pur di assumere $a_1 = 1 = a_2$:

$$a_n = \sum_{k=1}^{n-1} a_k a_{n-k};,$$

per ogni $n \geq 2$. (Suggerimento: dato k, $1 \leq k \leq (n-1)$, il numero a_n dei modi di disporre le parentesi sulla "parola"

$$X_1, X_2, \ldots, X_n$$

di n lettere è dato dal prodotto del numero di modi a_k di disporle sulle prime k lettere per quello a_{n-k} di disporle sulle rimanenti $(n-k)$). Più difficile è dimostrare la seguente espressione esatta:

$$a_n = \frac{1}{n}\binom{2n-2}{n-1}$$

e vedremo in seguito una tecnica per derivarla a partire dalla formula induttiva. Nell'esercizio 2.6.3 vedremo una interpretazione dei numeri di Catalano di interesse per l'informatica, poiché fornisce una interpretazione di questi numeri in termini della struttura delle memorie di massa ('hard disks').

1.14 Monoidi

Benchè un prodotto di insiemi sia essenzialmente determinato dalle funzioni che lo ammettono come codominio, il principale interesse di considerare un prodotto è dato dalle funzioni che lo ammettono come dominio. Una "*operazione n-aria*" su un insieme A è una funzione

$$f\colon A^n = A \times A \times \ldots \times A \longrightarrow A\,.$$

Dunque un'operazione 0-aria (o "*nullaria*") è una funzione $A^0 \longrightarrow A$, cioè semplicemente un elemento $a\colon * \longrightarrow A$ di A. Un'operazione 1-aria ("*unaria*") è un endomorfismo $A^1 = A \longrightarrow A$ di A. Nel caso di

un'operazione 2-aria ("*binaria*")

$$f\colon A^2 = A \times A \longrightarrow A$$

si usa indicare il valore $f(a,b)$ di f su una coppia $\langle a,b\rangle$ con la notazione

$$afb\,.$$

Esempi di operazioni binarie sono ben noti:

1. Se $\mathbf{N}$ indica l'insieme dei numeri naturali

$$\mathbf{N} = \{0,1,2,\ldots\ldots\}\,,$$

le operazioni di somma

$$+\colon \mathbf{N}\times\mathbf{N} \longrightarrow \mathbf{N}$$

e di prodotto

$$\cdot\,\colon \mathbf{N}\times\mathbf{N} \longrightarrow \mathbf{N}$$

sono esempi di operazioni binarie. Analogamente per le operazioni di somma e di prodotto sugli altri insiemi numerici noti, come l'insieme $\mathbf{Z}$ dei numeri interi relativi, l'insieme $\mathbf{Q}$ dei numeri razionali o l'insieme $\mathbf{R}$ dei numeri reali.

2. Se $M(n)$ indica l'insieme delle matrici quadrate di ordine n a coefficienti interi (o razionali, o reali), l'operazione di prodotto di matrici è un'operazione binaria

$$\cdot\colon M(n)\times M(n) \longrightarrow M(n)\,.$$

3. Per ogni insieme X, l'insieme $\mathrm{End}(X)$ degli endomorfismi di X è dotato in modo naturale di un'operazione binaria

$$\mathrm{End}(X)\times\mathrm{End}(X) \longrightarrow \mathrm{End}(X)$$

data dalla composizione di funzioni, cioè se f e g sono due endomorfismi di X definiamo il loro prodotto fg come la loro *composizione* (in quest'ordine!).

4. Per ogni insieme X, si possono considerare le n-uple ordinate $\langle x_1, x_2, \ldots, x_n\rangle$ di elementi di X, cioè gli elementi di X^n, come "parole" $x = x_1x_2\ldots x_n$ di lunghezza $n > 0$ sull'"alfabeto" X. È chiaro che data un'altra parola $y = y_1y_2\ldots y_m$ di lunghezza m si può considerare la parola $xy = x_1x_2\ldots x_ny_1y_2\ldots y_m$ e che ciò definisce una operazione binaria sull'insieme

$$W(X) = \sum_{n>0} X^n$$

di tutte le parole di tutte le possibili lunghezze, che chiameremo operazione di "concatenazione".

Se A è un insieme finito (con un numero abbastanza piccolo di elementi), può essere utile descrivere un'operazione binaria su A con una tabella. Ad esempio, se $\mathbf{2}$ indica l'insieme dei valori di verità "falso" e "vero", che possiamo convenire di indicare con 0 e 1 rispettivamente, le operazioni logiche di congiunzione ("e") e di disgiunzione ("o") sono operazioni binarie:

$$\wedge : \mathbf{2}\times\mathbf{2} \longrightarrow \mathbf{2} \quad , \quad \vee : \mathbf{2}\times\mathbf{2} \longrightarrow \mathbf{2}$$

che possono essere descritte mediante le tabelle:

$\wedge$	0	1
0	0	0
1	0	1

$\vee$	0	1
0	0	1
1	1	1

Un ulteriore esempio di operazione binaria su $\mathbf{2}$ è dato dalla nostra esperienza con i numeri *pari* e quelli *dispari*: se identifichiamo 0 con il concetto di numero pari e 1 con quello di numero dispari, allora la somma di numeri induce una somma su tali concetti, descritta dalla tabella

$+$	0	1
0	0	1
1	1	0

Tali tabelle vengono dette "*tavola*" della operazione che si sta considerando. Come ulteriore esempio scriviamo la tavola di moltiplicazione dell'operazione binaria $\mathrm{End}(\mathbf{2})\times\mathrm{End}(\mathbf{2}) \longrightarrow \mathrm{End}(\mathbf{2})$ data dalla

composizione di funzioni. Elenchiamo dapprima i quattro elementi di $\mathrm{End}(\mathbf{2})$:

$$\begin{array}{lll} f_1(0)=0 & , & f_1(1)=1 \\ f_2(0)=1 & , & f_2(1)=0 \\ f_3(0)=0 & , & f_3(1)=0 \\ f_4(0)=1 & , & f_4(1)=1 \end{array}$$

La tavola di moltiplicazione risulta essere la seguente:

$\circ$	f_1	f_2	f_3	f_4
f_1	f_1	f_2	f_3	f_4
f_2	f_2	f_1	f_4	f_3
f_3	f_3	f_3	f_3	f_3
f_4	f_4	f_4	f_4	f_4

Ad esempio, l'elemento $f_2 f_3$ si calcola così: $(f_2 f_3)(0) = f_2(f_3(0)) = f_2(0) = 1$; $(f_2 f_3)(1) = f_2(f_3(1)) = f_2(0) = 1$; dunque $f_2 f_3 = f_4$.

In generale le operazioni binarie che occorrono negli esempi concreti non sono del tutto arbitrarie, ma soddisfano delle *identità*, la più importante delle quali è l' "*associatività*". Un'operazione binaria $\circ\colon A \times A \longrightarrow S$ su A è detta "*associativa*" se, per ogni terna ordinata a, b, c di elementi di A, vale:

$$(a \circ b) \circ c = a \circ (b \circ c)\,.$$

Usando il principio di estensionalità per le funzioni, si vede che la precedente identità può essere espressa dicendo che le due possibili funzioni $A \times A \times A \longrightarrow A$ che si ottengono per composizione dal diagramma

$$\begin{array}{ccc} A \times A \times A & \xrightarrow{1_A \times \circ} & A \times A \\ {\scriptstyle \circ \times 1_A}\downarrow & & \downarrow{\scriptstyle \circ} \\ A \times A & \xrightarrow[\circ]{} & A\,, \end{array}$$

sono uguali o, in altri termini, dicendo che il precedente diagramma "*commuta*". Si osservi che se non si usasse la convenzione di scrivere

il risultato $\circ(a,b)$ di una operazione binaria come $a \circ b$, la associatività avrebbe la forma più complicata $\circ[\circ(a,b),c] = \circ[a,\circ(b,c)]$. La validità della legge associativa ci autorizza a non mettere parentesi quando scriviamo una composizione di tre elementi. Questo fatto si estende anche alle composizioni di più di tre elementi, ma una dimostrazione rigorosa di tale affermazione richiede qualche sforzo.

Definizione 1.14.1 . *Si dice "semigruppo" una coppia* $(A,\circ)$ *dove* A *è un insieme e "*$\circ$*" è una operazione binaria associativa su* A.

Può accadere che su uno stesso insieme siano definite diverse operazioni di semigruppo. Ad esempio, sull'insieme $\mathbf{N}$ dei numeri naturali sono definite due operazioni di semigruppo, la somma ed il prodotto. Anche sull'insieme $\mathbf{2}$ dei valori di verità sono definite almeno due operazioni di semigruppo, la congiunzione e la disgiunzione (se ne verifichi l'associatività). Tutte le operazioni binarie sopra considerate sono associative (esercizio).

Una ulteriore caratteristica delle operazioni binarie considerate fino ad ora è l'esistenza di un elemento particolare $e \in A$, tale che per ogni $a \in A$ si ha:

$$e \circ a = a = a \circ e\,.$$

Un tale elemento viene detto "*elemento neutro* (o "*unità*") per l'operazione "$\circ$". Negli esempi precedenti, l'elemento neutro in $(\mathbf{N},+)$ è il numero 0, in $(\mathbf{N},\cdot)$ è il numero 1, in $(\mathbf{2},\wedge)$ è il valore di verità "vero", mentre in $(\mathbf{2},\vee)$ è il valore di verità "falso", così come nell'esempio dei pari e dispari $(\mathbf{2},+)$. Ancora, nell'esempio $(M(n),\cdot)$ delle matrici quadrate di ordine n con operazione data dal prodotto di matrici è la matrice identica I_n e nel semigruppo $(\mathrm{End}(X),\cdot)$ è la funzione identità.

Si osservi che un ragionamento perfettamente simile a quello svolto in 1.1 sulla unicità della funzione identità, porta a concludere che un elemento neutro per un'operazione binaria, se esiste, è *unico*. Infatti, se esistesse un altro elemento, diciamo e', tale che $e' \circ a = a = a \circ e'$ per ogni $a \in A$, allora $e \circ e' = e$ perchè e' è un elemento neutro ed $e \circ e' = e'$ perchè e è elemento neutro.

Definizione 1.14.2 *Un "monoide" è un semigruppo con elemento neutro.*

Gli esempi di operazioni binarie considerati fino ad ora sono perciò tutti dei monoidi. Un esempio di semigruppo che non è un monoide è dato da un qualsiasi insieme X con almeno due elementi con l'operazione definita dalla (prima) proiezione, cioè $x \circ y = x$ (esercizio). Anche l'insieme $W(X)$ delle parole su X con l'operazione di concatenazione (detto "*semigruppo delle parole sull'alfabeto* X") è un semigruppo che non è un monoide; tuttavia, aggiungendo la "*parola vuota*" si ottiene un monoide che denoteremo con $(X)^*$ e chiameremo "*monoide delle parole sull'alfabeto* X".

Un ulteriore esempio di identità che si può considerare per una operazione binaria è la "*commutatività*". Un'operazione binaria "$\circ$" su un insieme A è detta "*commutativa*" se per ogni $\langle a, b\rangle \in A \times A$ si ha:

$$a \circ b = b \circ a\,.$$

Se $(A, \circ)$ è un semigruppo o un monoide per cui l'operazione è commutativa, diremo che $(A, \circ)$ è un semigruppo o un monoide *commutativo*, o anche *abeliano*. Si noti che in tal caso per assicurare che un elemento e è l'elemento neutro per l'operazione, è sufficiente la validità di una sola delle due identità $a \circ e = a = e \circ a$. Gli esempi precedenti $(\mathbf{N}, +)$, $(\mathbf{N}, \cdot)$, $(\mathbf{Q}, +)$ e tutti gli altri esempi "numerici" sono monoidi commutativi, così come $(\mathbf{2}, \wedge)$, $(\mathbf{2}, \vee)$ e $(\mathbf{2}, +)$, mentre i monoidi $(M(n), \cdot)$, $(\mathrm{End}(X), \cdot)$ *non* sono commutativi, a meno che $n = 1$ (si veda ad esempio la tavola di moltiplicazione di $(\mathrm{End}(\mathbf{2}), \cdot)$). Vedremo in seguito molti altri esempi di monoidi.

Le funzioni che normalmente si considerano tra monoidi sono quelle che "*conservano le operazioni*", nel senso della seguente

Definizione 1.14.3 . *Un "omomorfismo" o ("morfismo") di monoidi*

$$f\colon (A, \circ, e) \longrightarrow (B, \Box, e')$$

è una funzione $f\colon A \longrightarrow B$ *tale che:*

1) $f(a \circ b) = f(a) \Box f(b)$, *per ogni* $\langle a, b\rangle \in A \times A$

2) $f(e) = e'$.

Passiamo ora in rassegna qualche esempio. Molti altri ne vedremo in seguito.

1. Se $\mathbf{R}^{>0}$ denota l'insieme dei numeri reali positivi, allora $(\mathbf{R}^{>0}, \cdot, 1)$ è un monoide. La funzione esponenziale

$$\exp\colon \mathbf{R} \longrightarrow \mathbf{R}^{>0}$$

è un morfismo di monoidi

$$\exp\colon (\mathbf{R}, +, 0) \longrightarrow (\mathbf{R}^{>0}, \cdot, 1)\,,$$

poichè $\exp(a+b) = \exp(a)\exp(b)$, $\exp(0) = 1$.

2. La funzione $f\colon \mathbf{R} \longrightarrow M(2;\mathbf{R})$ definita da

$$f(x) = \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix}$$

è un morfismo di monoidi $f\colon (\mathbf{R}, +, 0) \longrightarrow (M(2;\mathbf{R}), \cdot, I_2)$ poichè $f(0) = I_2$ e $f(x+y) = f(x)f(y)$, come facilmente si verifica.

3. Se $(M(n;\mathbf{Q})\cdot, I_n)$ denota il monoide delle matrici quadrate di ordine n a coefficienti razionali e se $|\ \ |\colon M(n;\mathbf{Q}) \longrightarrow \mathbf{Q}$ denota la funzione "*determinante*", il *teorema di Binet* visto nel corso di Geometria può essere espresso dicendo che il determinante è un morfismo di monoidi

$$|\ \ |\colon (M(n;\mathbf{Q}), \cdot, I_n) \longrightarrow (\mathbf{Q}, \cdot, 1)\,.$$

4. Se $p \in \mathbf{Z}$ è un numero intero fissato, allora la funzione $f\colon \mathbf{Z} \longrightarrow \mathbf{Z}$ definita dalla moltiplicazione per p, cioè $f(x) = px$, è un morfismo $f\colon (\mathbf{Z}, +, 0) \longrightarrow (\mathbf{Z}, +, 0)$. Infatti, $f(0) = 0$ e $f(x+y) = p(x+y) = px + py = f(x) + f(y)$.

5. Se $(X)^*$ denota il monoide delle parole su un alfabeto X, la funzione $l\colon (X)^* \longrightarrow \mathbf{N}$ che ad ogni parola associa la sua lunghezza è un morfismo di monoidi da $(X)^*$ con l'operazione di concatenazione a $(\mathbf{N}, +, 0)$.

Diremo che un morfismo di monoidi $f\colon (A, \circ, e) \longrightarrow (B, \square, e')$ è iniettivo (risp. suriettivo, isomorfismo), se la funzione $f\colon A \longrightarrow B$ è iniettiva (risp. suriettiva, isomorfismo). Un fatto importante è il seguente:

Teorema 1.14.1 1) *Se*

$$f\colon (A,\circ,e) \longrightarrow (B,\Box,e')\,,\ g\colon (B,\Box,e') \longrightarrow (C,\diamond,e'')$$

sono morfismi di monoidi, allora la composizione

$$gf\colon (A,\circ,e) \longrightarrow (C,\diamond,e'')$$

è un morfismo di monoidi. Inoltre, per ogni monoide $(A,\circ,e)$ l'identità 1_A è un morfismo di monoidi e, se $(B,\Box,e')$ è un qualunque altro monoide, la funzione costante $A \longrightarrow B$ di valore e' è un morfismo di monoidi $(A,\circ,e) \longrightarrow (B,\Box,e')$.

2) *Se $f\colon (A,\circ,e) \longrightarrow (B,\Box,e')$ è un isomorfismo di monoidi, allora la funzione inversa $f^{-1}\colon B \longrightarrow A$ è ancora un morfismo di monoidi $f\colon (B,\Box,e') \longrightarrow (A,\circ,e')$.*

DIMOSTRAZIONE. Le dimostrazioni sono semplicissime. A titolo di esempio, dimostriamo la 2): da $f(e) = e'$ si ha $f^{-1}(e') = e$; inoltre, per dimostrare che $f^{-1}(x\Box y) = f^{-1}(x) \circ f^{-1}(y)$, poichè f è iniettiva basta dimostrare che i due elementi $f^{-1}(x\Box y)$ e $f^{-1}(x) \circ f^{-1}(y)$ assumono lo stesso valore nella funzione f:
$f[f^{-1}(x\Box y)] = x\Box y$;
$f[f^{-1}(x) \circ f^{-1}(y)] = f[f^{-1}(x)]\Box f[f^{-1}(y)]$, (perchè f è un morfismo) $= x\Box y$. ■
L'esempio del logaritmo è classico: poichè la funzione esponenziale è un morfismo di monoidi $\exp\colon (\mathbf{R},+,0) \longrightarrow (\mathbf{R}^{>0},\cdot,1)$, e poichè il logaritmo è definito come la sua funzione inversa, allora anche il logaritmo è un morfismo di monoidi $\log\colon (\mathbf{R}^{>0},\cdot,1) \longrightarrow (\mathbf{R},+,0)$ e dunque:

$$\log(xy) = \log(x) + \log(y), \quad \log(1) = 0\,.$$

1.15 Gruppi

L'esperienza con le funzioni ci conduce alla seguente

Definizione 1.15.1 *Sia $(M,\circ,e)$ un monoide. Un elemento $a \in M$ è detto "invertibile" se esiste un elemento $b \in M$ tale che:*

$$a \circ b = e = b \circ a\,,$$

essendo e l'elemento neutro del monoide.

Con un ragionamento del tutto simile a quello fatto per le funzioni invertibili (isomorfismi), si vede che se un tale elemento b esiste, esso è *unico*; cioè, se esistesse un altro elemento b' tale che $a \circ b' = e = b' \circ a$, allora $b = b'$. Infatti:
$b = e \circ b = (b' \circ a) \circ b = b' \circ (a \circ b) = b' \circ e = b'$.
Se a è invertibile, l'unico elemento b tale che $a \circ b = e = b \circ a$ viene chiamato "*inverso*" di a e viene denotato con a^{-1}. Nel caso che l'operazione binaria sia denotata con il simbolo "+", come ad esempio la somma nei monoidi "numerici", l'elemento neutro viene denotato con "0" e l'inverso di a con "$-a$" e viene chiamato "*opposto*" (notazione "additiva"). Si osservi anche qui che l'unicità dell'inverso implica che l'inverso di un elemento invertibile è a sua volta invertibile, poichè

$$(a^{-1})^{-1} = a$$

e che la composizione di due elementi invertibili è invertibile, poichè

$$(a \circ b)^{-1} = b^{-1} \circ a^{-1}\,,$$

con un argomento perfettamente simile a quello che abbiamo dato per le funzioni invertibili in 1.2.

Definizione 1.15.2 *Un "gruppo" è un monoide in cui ogni elemento è invertibile.*

Una osservazione importante è che l'unicità dell'inverso permette di definire una operazione unaria

$$(\ \ )^{-1}\colon M \longrightarrow M$$

data proprio dall'inverso di ogni elemento di un gruppo, e che quindi la nozione di gruppo può essere data in modo equivalente come un insieme $(M, \circ, (\ \ )^{-1}, e)$ munito di tre operazioni, una binaria, una unaria e una nullaria, che soddisfano le identità definenti un monoide, più quelle che qualificano x^{-1} come l'inverso di x. In altre parole, anche un gruppo è una *struttura algebrica*, cioè un insieme con operazioni che soddisfano delle identità. Passiamo in rassegna gli esempi precedenti:

1. I monoidi $(\mathbf{Z}, +)$, $(\mathbf{Q}, +)$, $(\mathbf{R}, +)$ sono tutti gruppi (commutativi).

2. Negli esempi di operazioni di monoide su $\mathbf{2}$, l'unica operazione di gruppo è quella di somma, nell'interpretazione di 0 e 1 come pari e dispari.

3. I monoidi $(\mathbf{Z}, \cdot)$, $(\mathbf{Q}, \cdot)$, $(\mathbf{R}, \cdot)$ *non* sono gruppi. Negli ultimi due l'unico motivo è che il numero 0 non ha un inverso. Ma se lo togliamo e indichiamo con $\mathbf{Q}^*$ e $\mathbf{R}^*$ gli insiemi $\mathbf{Q}$ e $\mathbf{R}$ cui è stato tolto il numero 0, allora $(\mathbf{Q}^*, \cdot)$ e $(\mathbf{R}^*, \cdot)$ *sono* gruppi.

4. Se $(M, \circ)$ è un monoide, indichiamo con $(M^*, \circ)$ il monoide degli elementi invertibili di M rispetto all'operazione $\circ$. Rispetto a tale operazione, M^* è un gruppo, per le proprietà sopra descritte degli elementi invertibili. A esempio, nel caso del monoide $(M(n; \mathbf{Q}), \cdot)$ delle matrici quadrate di ordine n a coefficienti nei numeri razionali, il gruppo degli elementi invertibili è il gruppo $\mathrm{GL}(n; \mathbf{Q})$ delle matrici quadrate di ordine n a coefficienti razionali il cui determinante è diverso da zero (si ricordino i teoremi di Binet e di Cramer studiati nel corso di Geometria). Ancora, se X è un insieme qualsiasi, gli elementi invertibili del monoide $\mathrm{End}(X)$ degli endomorfismi di X rispetto all'operazione definita dalla composizione di funzioni è un gruppo, che denotiamo con $\mathrm{Aut}(X)$ e che chiamiamo gruppo degli automorfismi di X. Se X è un insieme finito con n elementi, allora tale gruppo è denotato con S_n ed è chiamato anche "*gruppo simmetrico*" o anche "*gruppo delle permutazioni su n elementi*" ed ha cardinalità (o ordine) $n!$. Studieremo in seguito più approfonditamente tali gruppi. A titolo di esempio, se X ha due elementi, si provi che S_2 ha due elementi e che, se li indichiamo con f_1 e f_2, la tavola di moltiplicazione è la seguente:

$\circ$	f_1	f_2
f_1	f_1	f_2
f_2	f_2	f_1

Non è difficile convincersi che cambiando f_1 in 0 e f_2 in 1 tale tavola coincide con quella di $(\mathbf{2}, +)$.

Un'ultima utile osservazione sulla costruzione del gruppo degli elementi invertibili di un monoide è che essa si estende agli omomorfismi: se $f\colon (M, \circ, e) \longrightarrow (N, \square, u)$ è un omomorfismo di monoidi,

allora f induce un omomorfismo

$$f\colon (M^*, \circ, e) \longrightarrow (N^*, \square, u)\,,$$

poichè f *conserva gli inversi*: se $x \circ x^{-1} = e = x^{-1} \circ x$, allora

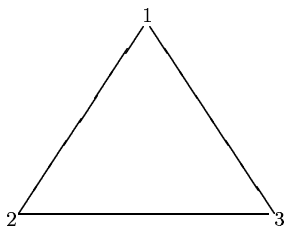
$f(x) \square f(x^{-1}) = f(x^{-1} \circ x) = f(e) = u$

e similmente

$f(x^{-1}) \square f(x) = f(x \circ x^{-1}) = f(e) = u$;

dunque $f(x^{-1}) = f(x)^{-1}$.

5. Come ulteriore esempio si può considerare S_3 e determinare la tavola della sua operazione. Gli elementi di S_3 sono $3! = 6$ e per determinarli, così come per determinare la tavola dell'operazione può essere utile osservare che S_3 ha la seguente interpretazione geometrica. Consideriamo un triangolo equilatero:



e consideriamo tutte le sue possibili simmetrie, cioè le rotazioni di 120 gradi intorno al suo baricentro e i ribaltamenti rispetto ad una altezza (= mediana = bisettrice). Ad esempio, se scriviamo:

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

intendiamo che operando il ribaltamento rispetto all'altezza relativa al vertice 2, il risultato è che il vertice 1 va nel vertice 3, il vertice 3 va nel vertice 1 e il vertice 2 rimane fisso. Ogni elemento di S_3 può essere interpretato in questo modo e, viceversa, ogni simmetria può essere interpretata come un elemento di S_3; inoltre, il prodotto di due elementi di S_3 può essere pensato come operare le due simmetrie corrispondenti "una dopo l'altra". Formalmente, se indichiamo con Δ_3 l'insieme delle simmetrie del

triangolo equilatero (sono 6) si può verificare che l'applicazione di due simmetrie "una dopo l'altra" definisce un'operazione binaria di gruppo su Δ_3 e che l'interpretare ogni elemento di Δ_3 in una permutazione di S_3 definisce una funzione $f\colon \Delta_3 \longrightarrow S_3$ che è un *isomorfismo* di gruppi.

6. Naturalmente si può procedere con lo stesso tipo di esempio per $n = 4$, cioè si può considerare il gruppo Δ_4 delle simmetrie di un quadrato, che sono costituite dalle rotazioni dei quattro multipli di 90 gradi rispetto al baricentro e dai ribaltamenti rispetto alle diagonali e agli assi e quindi da 8 elementi. Numerando i vertici del quadrato, si può definire un omomorfismo iniettivo $\Delta_4 \longrightarrow S_4$, che *non* è un isomorfismo (S_4 ha 24 elementi), ecc. . È chiaro come tutto ciò si generalizza ai poligoni *regolari* con n lati. Indicheremo con Δ_n il gruppo delle simmetrie di un poligono regolare con n lati e chiameremo "*gruppi diedrici*" tali gruppi. Ricordiamo che fissato un vertice del poligono regolare di n lati, se chiamiamo R la rotazione di $1/n$ di angolo giro e D il ribaltamento intorno all'asse passante per quel vertice, allora $R^n = 1$ (dove 1 indica la simmetria identità che è la rotazione di 0 gradi e R^i indica la rotazione di $1/n$ di angolo giro eseguita i volte) e $D_2 = 1$; inoltre, si può constatare che i $2n$ elementi dati dalle composizioni di simmetrie:

   $R^0 = 1, R, R^2, \ldots, R^{n-1}, D, RD, R^2D, \ldots, R^{n-1}D,$

   esauriscono *tutte* la simmetrie del poligono regolare con n lati, cioè sono un elenco completo degli elementi di Δ_n, perchè un semplice ragionamento geometrico porta a concludere che *tutti* i ribaltamenti si ottengono da uno dato D seguito da una rotazione. Si esprime questo fatto dicendo che D e R sono un insieme di *generatori* per il gruppo Δ_n. Infine, non è difficile convincersi che la simmetria DR è uguale alla $R^{n-1}D$; infatti entrambi tali prodotti hanno come effetto di invertire l'ordine dei vertici e di portare il vertice 1 nel vertice n. L'osservazione importante è che *tutti i prodotti degli elementi di Δ_n si possono calcolare conoscendo solo queste tre relazioni fondamentali ("relazioni di definizione")*:

$$R^n = 1, \quad D^2 = 1, \quad DR = R^{n-1}D\,.$$

Infatti, siano R^jD^i e R^kD^l due elementi di Δ_n (dove j e k sono indici che variano tra 0 e $n-1$, mentre i e l sono indici che variano tra 0 e 1); per determinare il loro prodotto, si può usare ripetutamente la terza delle relazioni di definizione per spostare tutte le occorrenze di R a sinistra e quelle di D a destra; a questo punto si usano le altre due relazioni di definizione per ridurre ciò che si ottiene ad uno degli elementi di Δ_n. Ad esempio in Δ_4 il prodotto $(RD)(R^2D)$ può essere calcolato nel modo seguente:

$RDR^2D = R(DR)RD = R(R^3D)RD = R^4DRD = DRD = R^3DD = R^3$.

In tal modo si può determinare completamente la tavola di moltiplicazione di Δ_4 (esercizio) e constatare, ad esempio, che Δ_4 *non* è commutativo.

Concludiamo con una utile osservazione sugli omomorfismi di gruppi. In principio, un omomorfismo di gruppi è lo stesso che un omomorfismo di monoidi, con in più la proprietà di conservare anche l'inverso. Tuttavia può accadere che se i monoidi hanno particolari proprietà, allora per un morfismo di monoidi la condizione di conservare l'elemento neutro è automaticamente soddisfatta, ad esempio, se il monoide codominio $(B, \Box, e')$ ha la "*proprietà di cancellazione (a sinistra)*":
per ogni $a, b, c \in B$*: se* $a\Box b = a\Box c$*, allora* $b = c$.
Esempi di tali monoidi sono i monoidi $(\mathbf{N}, +, 0)$ e $(\mathbf{N}^{>0}, \cdot, 1)$, dove $\mathbf{N}^{>0}$ indica l'insieme dei naturali diversi da 0. Una classe importante di esempi è costituita dai *gruppi*. Infatti, dato che per ogni elemento esiste l'inverso, se $ab = ac$, allora $a^{-1}ab = a^{-1}ac$, dunque $b = c$.

Teorema 1.15.1 *Siano* $(A, \circ, e)$ *e* $(B, \Box, e')$ *due monoidi. Se* $(B, \circ, e')$ *ha la proprietà di cancellazione, allora ogni funzione* $f: A \longrightarrow B$ *che conserva l'operazione binaria conserva anche l'elemento neutro, cioè è un morfismo di monoidi.*

In particolare, se sono gruppi, allora ogni funzione $f: A \longrightarrow B$ *che conserva l'operazione binaria è un morfismo di gruppi, perchè conserva anche l'inverso, cioè* $f(x^{-1}) = f(x)^{-1}$.

DIMOSTRAZIONE. Poichè $e = e \circ e$, si ha $f(e) = f(e \circ e) = f(e)\Box f(e)$; quindi $f(e)\Box e' = f(e) = f(e)\Box f(e)$ e per la proprietà di cancellazione si ottiene $e' = f(e)$.

In particolare, se i due monoidi sono gruppi, se $f: A \longrightarrow B$ conserva l'operazione binaria, allora è un morfismo di gruppi, perchè conserva anche l'inverso; infatti:
$f(x)\Box f(x^{-1}) = f(x \circ x^{-1}) = f(e) = e'$;
e similmente $f(x^{-1})\Box f(x) = e'$; dunque $f(x^{-1})$ è l'inverso di $f(x)$, cioè $f(x^{-1}) = f(x)^{-1}$. ■ Ad esempio, osservando che i reali positivi sono un gruppo rispetto al prodotto, applicando questo teorema si ha che le sole proprietà $\exp(x+y) = \exp(x)\exp(y)$, $\log(xy) = \log(x) + \log(y)$ della funzione esponenziale e del logaritmo implicano che valgono anche le proprietà

$$\exp(-x) = \frac{1}{\exp(x)} \quad , \quad \exp(0) = 1$$

$$\log(\frac{1}{x}) = -\log(x) \quad , \quad \log(1) = 0\,.$$

1.16 Anelli

Fino ad ora abbiamo considerato strutture costituite da un insieme munito di una singola operazione binaria, ma l'esperienza con i numeri naturali e con gli altri insiemi numerici che abbiamo fin dai primi anni di scuola, ci conduce a considerare insiemi muniti di due operazioni binarie

Definizione 1.16.1 *Un "anello" è un insieme A munito di due operazioni binarie "somma" e "prodotto"*

$$+: A \times A \longrightarrow A \quad , \quad \cdot : A \times A \longrightarrow A$$

che denoteremo come di consueto con $a+b$ e ab rispettivamente, che soddisfano le proprietà:

1) *$(A, +)$ è un gruppo commutativo, il cui elemento neutro è denotato con 0;*

2) *$(A, \cdot)$ è un monoide, il cui elemento neutro è denotato con 1;*

3) *valgono le proprietà distributive:*

$$(x+y)z = xz + yz \quad , \quad z(x+y) = zx + zy\,,$$

per ogni $x, y, z \in A$.

L'anello viene detto "commutativo", se il prodotto è commutativo.

Esempi di anelli sono già noti a tutti fin dai primi anni degli studi: gli anelli commutativi dei numeri interi $(\mathbf{Z}, +, \cdot)$, dei numeri razionali $(\mathbf{Q}, +, \cdot)$, eccetera. Un nuovo esempio fondamentale di anello (non commutativo) è quello $M(n; A)$ delle matrici a quadrate di ordine n a coefficienti in un anello (commutativo) $(A, +, \cdot)$, introdotto nel corso di Geometria. Ricordiamo anche che un esempio ben noto è quello dell'*anello $A[x]$ dei polinomi a coefficienti in un anello commutativo* $(A, +, \cdot)$: i suoi elementi sono i "*polinomi*", che sono scritture formali

$$p = p(x) = \sum_{i=0}^{n} p_i x^i$$

(n viene detto grado del polinomio) e, osservando che un polinomio è conosciuto quando sono specificati i coefficienti p_i, somma e prodotto sono definiti dalle seguenti regole: se

$$q = q(x) = \sum_{i=0}^{m} q_i x^i$$

è un altro polinomio, allora i coefficienti $(p+q)_i$ e $(pq)_i$ della somma $p+q$ e del prodotto pq sono dati dalle regole

$$(p+q)_i = p_i + q_i$$

$$(pq)_i = \sum_{h+k=i} p_h q_k = \sum_{h=0}^{i} p_h q_{i-h} \, .$$

Il lettore dovrebbe verificare che le regole descritte corrispondono a quelle che già conosce e che tali definizioni soddisfano gli assiomi di anello commutativo. Inoltre, indicando con $gr(p)$ il grado di un polinomio p, si vede facilmente che $gr(p+q) \leq max[gr(p), gr(q)]$ e $gr(pq) \leq gr(p) + gr(q)$ e che in questa ultima disuguaglianza vale il segno di uguaglianza se e solo se in A non esistono *"divisori di zero"*, cioè coppie di elementi non nulli a, b tali che $ab = 0$. Tali anelli si chiamano *"domini di integrità"*; vedremo in seguito esempi anelli commutativi che *non* sono domini di integrità (2.8).

Tutte le regole del calcolo imparate fino ad ora sono conseguenze degli assiomi di anello. Ad esempio, il "raccogliere a fattore comune" altro non è che l'applicazione della proprietà distributiva. Un'altra importante conseguenza della distributività è la "*regola dei segni*". Vediamo come si dimostra:

Lemma 1.16.1 *In ogni anello* $(A, +, \cdot)$ *si ha:*

1) $a0 = 0a = 0$, *per ogni* $a \in A$.

2) $(-a)b = a(-b) = -(ab)$, *per ogni* $a, b \in A$; *in particolare,* $(-1)(-1) = -(-1) = 1$ *(regola dei segni).*

3) *Se* $0 = 1$, *allora* A *ha un solo elemento.*

DIMOSTRAZIONE. 1) $a0 = a(0+0) = a0 + a0$, da cui sommando $-(a0)$ a entrambi i membri si ottiene $a0 = 0$. Similmente si prova $0a = 0$.

2) $0 = 0b = [a + (-a)]b = ab + (-a)b$, da cui sommando $-(ab)$ a entrambi i membri si ottiene $-(ab) = (-a)b$. Similmente si prova $-(ab) = a(-b)$.

3) Per ogni $a \in A$ si ha: $a = a1 = a0 = 0$, dunque A ha un solo elemento. ■

Un "*omomorfismo di anelli*" $f\colon (A, +, \cdot) \longrightarrow (B, +, \cdot)$ è una funzione $f\colon A \longrightarrow B$ che è un morfismo sia per la struttura di gruppo, sia per quella di monoide. Ricordando l'osservazione finale del paragrafo precedente e traducendola in notazione additiva, ciò equivale a richiedere le tre proprietà $f(x+y) = f(x)+f(y)$, $f(xy) = f(x)f(y)$ e $f(1) = 1$; per quanto ricordato, dalla prima di queste tre proprietà si ha che valgono anche $f(0) = 0$ e $f(-x) = -f(x)$.

1.17 Esercizi

1. Si provi che la commutatività della somma è una conseguenza delle altre proprietà, in particolare delle distributività. (*Suggerimento: si sviluppi nei due modi possibili il prodotto* $(1+1)(x+y)$.)

2. Sia $(A, +, \cdot)$ un anello commutativo, e sia $A[[x]]$ l'insieme delle *serie formali*, cioè delle scritture formali del tipo

$$a = \sum_{i=0}^{\infty} a_i x^i \,.$$

È chiaro che un tale oggetto è completamente individuato dalla *successione* $a\colon \mathbf{N} \longrightarrow A$ dei suoi coefficienti. Si provi che rispetto alla somma e al prodotto definiti esattamente come nel caso dei polinomi si ottiene un anello commutativo $A[[x]]$.

3. Ricordando che in un anello gli elementi invertibili sono quelli che sono invertibili rispetto alla struttura di monoide moltiplicativo, si provi che in un anello di polinomi gli unici invertibili sono i polinomi di grado 0 che sono invertibili nell'anello. Chi sono gli elementi invertibili in un anello di serie formali?

4. Se $(A, +, \cdot)$ è un anello commutativo, si provi che per ogni $a \in A$ la funzione

$$\mathrm{val}_a\colon A[x] \longrightarrow A$$

data dalla *valutazione in* a, cioè $\mathrm{val}_a p(x) = p(a)$, è un omomorfismo di anelli.

5. Si provi che una anello commutativo A è un dominio di integrità se e solo se vale la seguente "*legge di cancellazione*":

$$ab = ac \text{ e } a \neq 0 \quad \Rightarrow \quad b = c\,.$$

6. Si provi che sull'anello dei polinomi $A[x]$ a coefficienti in una anello A esiste un'unica "*derivazione*", cioè una funzione

$$D\colon A[x] \longrightarrow A[x]$$

tale che

   (a) $D[p(x) + q(x)] = D[p(x)] + D[q(x)]$, $D[ap(x)] = aD[p(x)]$;
   (b) $D[p(x)q(x)] = p(x)D[q(x)] + D[p(x)]q(x)$;
   (c) $D(x) = 1$, $D(1) = 0$.

Esiste una tale funzione sull'anello $A[[x]]$ delle serie formali a coefficienti in A?

7. In un anello commutativo $(A, +, \cdot)$ in cui ogni somma di 1 sia invertibile, si consideri la serie formale

$$\exp(x) = \sum_{i=0}^{\infty} \frac{x^i}{i!} .$$

si provi

$$D[\exp(x)] = \exp(x) .$$

8. Si provi che il polinomio $(1 - x)$, considerato come serie formale a coefficienti in un anello commutativo $(A, +, \cdot)$, è invertibile e che la serie formale inversa è

$$\frac{1}{1-x} = \sum_{i=0}^{\infty} x^i .$$

9. Sia $(A, +)$ un gruppo *commutativo* e si denoti con $\mathrm{End}_+(A)$ l'insieme degli omomorfismi di gruppo $f: A \longrightarrow A$. Si provi che definendo la somma di due elementi $f, g \in \mathrm{End}_+(A)$ mediante

$$(f + g)(a) =_{df} f(a) + g(a)$$

(somma "puntuale") si ottiene ancora un omomorfismo $A \longrightarrow A$. Si provi che rispetto a tale somma e rispetto al prodotto definito mediante la composizione, l'insieme $\mathrm{End}_+(A)$ è un anello, non necessariamente commutativo.

Considerando invece l'insieme $\mathrm{End}(A)$ di *tutte* le funzioni $A \longrightarrow A$ con le stesse operazioni, si provi che si ottiene una struttura che soddisfa tutti gli assiomi di anello, esclusa una delle due distributività ("*quasianello*").

1.18 Prodotti e sottostrutture

Ci sono tre costruzioni fondamentali che permettono di generare nuove strutture algebriche (monoidi, gruppi, anelli, ecc.) a partire da strutture date. Esse sono il "*prodotto di strutture*", le "*sottostrutture*" e le

"*strutture quoziente*". Illustriamo ora le prime due, mentre la terza, che è un po' più delicata, la affronteremo nel capitolo 3.

Se $(A, \circ)$ e $(B, \Box)$ sono due semigruppi, allora definiamo il semigruppo prodotto $(A, \circ) \times (B, \Box)$ nel modo seguente: consideriamo l'insieme $A \times B$ prodotto degli insiemi A e B e definiamo l'operazione binaria su $A \times B$ come il prodotto "*puntuale*", cioè

$$\langle a, b\rangle \cdot \langle a', b'\rangle = \langle a \circ a', b \Box b'\rangle \,.$$

In termini diagrammatici, l'operazione "$\cdot$" può essere descritta nel modo seguente:

$$A \times B \times A \times B \xrightarrow{1_A \times \sigma \times 1_B} A \times A \times B \times B \xrightarrow{\circ \times \Box} A \times B \,,$$

dove $\sigma\colon B \times A \longrightarrow A \times B$ indica l'isomorfismo "simmetria" (si veda il paragrafo 1.16). È immediato constatare che tutte le proprietà equazionali che valgono per entrambe le operazioni continuano a valere per l'operazione puntuale sul prodotto. Dunque, se $(A, \circ)$ e $(B, \Box)$ sono semigruppi, anche il prodotto puntuale è un semigruppo e, se entrambi sono commutativi, anche l'operazione puntuale è commutativa. Inoltre, se $(A, \circ)$ e $(B, \Box)$ sono monoidi o gruppi, allora anche il prodotto puntuale è un monoide o un gruppo, l'elemento neutro essendo definito dalla coppia degli elementi neutri e inverso dalla coppia degli inversi. Ciò si estende anche al caso in cui ci sia più di una operazione binaria: ad esempio, se si parte da anelli, sul prodotto si possono definire ancora due operazioni puntuali e si può facilmente vedere che soddisfano ancora gli assiomi di anello.

In particolare si può considerare il prodotto $A \times A = A^2$ di una struttura con se stessa e quindi anche i prodotti iterati (potenze) A^n, per $n \geq 2$. Ad esempio, se consideriamo il gruppo commutativo $(\mathbf{R}, +, 0)$ dei numeri reali, allora l'operazione puntuale su $\mathbf{R}^n$ definisce la somma di vettori a n componenti studiata nel corso di Geometria, in cui appunto la somma definita "componente per componente". Naturalmente, si può definire anche il prodotto puntuale, ottenendo così un anello commutativo.

Per il prodotto di strutture dello stesso tipo (semigruppi, monoidi, gruppi, eccetera) vale la stessa proprietà universale che abbiamo visto

per il prodotto di insiemi, rispetto a tutte le altre strutture dello stesso tipo ed ai loro omomorfismi:
le proiezioni sono omomorfismi $p_A: A \times B \longrightarrow A$ e $p_B: A \times B \longrightarrow B$ e, se X è una qualsiasi altra struttura dello stesso tipo e $f: X \longrightarrow A$, $g: X \longrightarrow B$ sono due omomorfismi, allora l'unica funzione nel prodotto $\langle f, g\rangle: X \longrightarrow A \times B$ è un omomorfismo.

Una semplice ma importante osservazione riguarda la commutatività: se $(A, \circ)$ è un monoide, l'operazione stessa è una funzione $A \times A \longrightarrow A$ e, dato che $A \times A$ è un monoide per l'operazione puntuale, possiamo chiederci se essa stessa è un omomorfismo dal monoide $A \times A$ con l'operazione puntuale al monoide A stesso. Il lettore è invitato a verificare che la risposta è affermativa se e solo se l'operazione è *commutativa*.

Un altro importante concetto che si riferisce solo alle strutture e non alle identità che in esse possono essere soddisfatte è quello di *sottostruttura* di una struttura.

Definizione 1.18.1 *Se $(A, \circ)$ ùn insieme munito di una operazione binaria, un sottoinsieme $U \subseteq A$ è una sottostruttura se vale la condizione: per ogni $a, b \in U$, $a \circ b \in U$.*

È chiaro come questa definizione si estende a strutture definite anche da altre operazioni, ad esempio nullarie o unarie. Ad esempio, se $(A, \circ, e)$ è un insieme munito di una operazione binaria e di una nullaria (cioè di un elemento $e \in A$), allora un sottoinsieme $U \subseteq A$ è una sottostruttura se vale l'ulteriore condizione $e \in U$. Se inoltre in A è anche assegnata una operazione unaria $f: A \longrightarrow A$ (come l'inverso nei gruppi), allora il sottoinsieme U per essere una sottostruttura deve essere chiuso anche rispetto a tale operazione, cioè se $x \in U$, allora $f(x) \in U$.

Ne segue che se U è una sottostruttura di un semigruppo o di un monoide o di un gruppo, allora U stesso con le operazioni definite per restrizione di quelle su A è ancora una struttura dello stesso tipo per cui tutte le identità che valevano in A continuano a valere e dunque U per tali operazioni è ancora un semigruppo o un monoide o un gruppo. Il semigruppo così definito da una sottostruttura viene detto "*sottosemigruppo*" di $(A, \circ)$ e nel caso dei monoidi e dei gruppi viene detto "*sottomonoide*" e "*sottogruppo*" rispettivamente. È chiaro anche come

tutto ciò si estende al caso di strutture definite da più di una operazione binaria, come gli anelli.

Passiamo in rassegna alcuni esempi:

1. $\mathbf{N} \subseteq \mathbf{Z}$ *non* è sottostruttura della struttura di gruppo su $\mathbf{Z}$ data dalla somma, l'opposto e lo zero ($\mathbf{N}$ non è chiuso rispetto all'opposto). $\mathbf{N}$ è sottostruttura della struttura di monoide di $\mathbf{Z}$ data dalla somma e lo zero ed è anche sottostruttura della struttura di monoide di $\mathbf{Z}$ data dal prodotto e 1, cioè $\mathbf{N}$ è sottomonoide di $(\mathbf{Z}, +, 0)$ e di $(\mathbf{Z}, \cdot, 1)$. Similmente, $\mathbf{Z}$ è un sottogruppo del gruppo $(\mathbf{Q}, +, -, 0)$ dei razionali, mentre è un sottomonoide del gruppo $(\mathbf{Q}^*, , (\ )^{-1}, 1)$ dei numeri razionali non nulli. In altre parole, $\mathbf{Z}$ è un sottoanello dell'anello dei razionali. Ancora, l'insieme $\mathbf{Q}^{>0}$ dei numeri razionali positivi è un sottogruppo del gruppo $(\mathbf{Q}^*, \cdot, (\ )^{-1}, 1)$, mentre è un sottosemigruppo del gruppo $(\mathbf{Q}, +, -, 0)$, ecc. .

2. Il sottoinsieme $\{0, 1\} \subseteq \mathbf{N}$ è una sottostruttura del monoide $(\mathbf{N}, \cdot, 1)$, ma non del monoide $(\mathbf{N}, +, 0)$. Si osservi che la struttura di monoide di $\{0, 1\}$ come sottomonoide di $(\mathbf{N}, \cdot, 1)$ coincide con la struttura di monoide di $\{0, 1\}$ rispetto alla congiunzione logica quando pensiamo $\{0, 1\}$ come l'insieme $\mathbf{2}$ dei valori di verità.

3. Se $(M, \circ, e)$ è un monoide, l'insieme

$$M^* = \{m \in M \mid m \text{ invertibile}\} \subseteq M$$

è un sottomonoide di $(M, \circ, e)$, perchè il prodotto di due elementi invertibili è ancora un elemento invertibile e perchè l'elemento neutro è sempre invertibile. In particolare, M^* è un gruppo; tuttavia non possiamo dire che M^* sia un sottogruppo di $(M, \circ, e)$, non essendo $(M^*, \circ, e)$ stesso un gruppo; anzi, $(M, \circ, e)$ è un gruppo se e solo se $M^* = M$. Ricordiamo che se $(M, \circ, e)$ è il monoide delle matrici quadrate a coefficienti reali o razionali, allora M^* è il gruppo delle matrici invertibili, che per il teorema di Cramer, coincide con il gruppo delle matrici a determinante non nullo. Similmente, se $(M, \circ, e)$ è il monoide delle endofunzioni di un insieme X, allora M^* è il gruppo delle endofunzioni invertibili di X, cioè il gruppo delle permutazioni di X.

4. Se A è un insieme dotato di una struttura che contiene operazioni nullarie, allora l'insieme vuoto $\emptyset \subseteq A$ *non* può essere sottostruttura di A. Anzi, l'insieme vuoto $\emptyset$ è sottostruttura di A se e solo se la struttura di A non ha operazioni nullarie. In particolare, un sottomonoide di un monoide non può mai essere vuoto, dovendo contenere almeno l'elemento neutro. Per ogni monoide $(A, \circ, e)$, il sottoinsieme $\{e\}$ costituito dal solo elemento neutro è sempre un suo sottomonoide, che è anche un gruppo.

In certi casi si ha un fenomeno simile a quello visto per i morfismi. Cioè può accadere che particolari proprietà della struttura sull'insieme A possono ridurre il numero delle condizioni perchè un sottoinsieme $U \subseteq A$ sia una sottostruttura di quella di A. Ad esempio:

Teorema 1.18.1 *Se $(A, \cdot, (\ )^{-1}, e)$ è un gruppo finito (cioè A è un insieme finito), allora perchè un sottoinsieme $U \subseteq A$ sia un suo sottogruppo basta che U sia chiuso rispetto all'operazione binaria e sia non vuoto.*

DIMOSTRAZIONE. Dobbiamo verificare che U contiene l'elemento neutro e che è chiuso rispetto all'inverso. Poichè U non è vuoto, sia $a \in U$; consideriamo gli elementi: $a, a^2, \ldots a^n, \ldots$. Tali elementi appartengono tutti a U, perchè U è chiuso rispetto all'operazione binaria, e non possono essere tutti distinti, perchè A, quindi anche U, è finito. Dunque devono esistere due interi n e m, con $n \neq m$, tali che $a^n = a^m$; allora, se $n > m$, applicando l'operazione $n - m$ volte con l'inverso di a si ha che $a^{n-m} = e \in U$, perchè ogni potenza positiva di a è in U. Dunque, quando $n > m+1$, anche $a^{n-m-1} \in U$ ed è l'inverso di a; se $n = m+1$, allora $a = e$, che coincide con il suo inverso. Similmente se $m > n$. ■

Per finire, mostriamo che ad ogni omomorfismo $f\colon (A, \circ, e) \longrightarrow (B, \square, u)$ di monoidi sono associati canonicamente due sottostrutture, il "*nucleo*" e la "*immagine*" di f. Il nucleo $\ker(f)$ di f è l'immagine inversa dell'elemento neutro u, cioè

$$\ker(f) = \{a \in A \,|\, f(a) = u\} \subseteq A\,,$$

mentre l'immagine è l'immagine insiemistica di f, cioè

$$\mathrm{Im}(f) = \{b \in B \,|\, \text{esiste } a \in A \text{ tale che } f(a) = b\} \subseteq B\,.$$

Non è difficile mostrare che sia l'immagine sia il nucleo sono sottostrutture e quindi sono sottomonoidi di B e di A rispettivamente: se $a_1, a_2 \in \ker(f)$, cioè se $f(a_1) = u = f(a_2)$, allora, poichè f è un omomorfismo, $f(a_1 \circ a_2) = f(a_1) \Box f(a_2) = u \Box u = u$ e dunque $a_1 \circ a_2 \in \ker(f)$; sempre perchè f è un omomorfismo si ha anche $e \in \ker(f)$. Per quanto riguarda l'immagine, se $b_1, b_2 \in \mathrm{Im}(f)$, cioè se esistono $a, a_2 \in A$ tali che $f(a_1) = b_1$ e $f(a_2) = b_2$, allora, poichè f è un omomorfismo, $f(a_1 \circ a_2) = f(a_1) \Box f(a_2) = b_1 \Box b_2$; dunque esiste $a_3 = a_1 \circ a_2$ tale che $f(a_3) = b_1 \Box b_2$ e perciò $b_1 \Box b_2 \in \mathrm{Im}(f)$; sempre perchè f è un omomorfismo si ha anche $e \in \mathrm{Im}(f)$. Si osservi che nè per definire l'immagine, nè per mostrare che l'immagine è sottostruttura si usano gli elementi neutri; dunque l'immagine è una sottostruttura anche per i semigruppi.

Un uso frequente del nucleo è per mostrare che un dato omomorfismo è una funzione iniettiva: se $(A, \circ, e)$ e $(B, \Box, u)$ sono monoidi e f è un omomorfismo $f: A \longrightarrow B$, che come funzione è iniettiva, allora $\ker(f)$ è costituito dal solo elemento neutro e: infatti, se f è iniettiva e x è tale che $f(x) = e = f(e)$, allora necessariamente $x = e$. Tuttavia, se i due monoidi sono in realtà *gruppi*, allora vale anche il viceversa:

se ker(f) è costituito dal solo elemento neutro, allora l'omomorfismo f è una funzione iniettiva.

Infatti, se $f(a_1) = f(a_2)$, allora $f(a_1) \Box f(a_2)^{-1} = u$; poichè f è un omomorfismo, allora $f(a_1) \Box f(a_2)^{-1} = f(a_1 \circ a_2^{-1}) = u$ e dunque $a_1 \circ a_2^{-1} \in \ker(f)$; poichè $\ker(f)$ è costituito dal solo elemento neutro, si ha $a_1 \circ a_2^{-1} = e$, cioè $a_1 = a_2$.

Capitolo 2

Numeri

2.1 I Numeri Naturali

Fino ad ora abbiamo usato i numeri naturali ed i numeri interi basandoci sulla conoscenza più o meno intuitiva che ognuno ne ha. Ma per poter fare della matematica, cioè per poter fare dei calcoli esatti e per poter *dimostrare* le proprietà che ci occorrono per i calcoli che dobbiamo fare, è necessario poter esprimere in modo formale ciò che intendiamo per numeri naturali; e questo in matematica significa esprimere con un *sistema di assiomi* ciò che si intende per l'insieme $\mathbf{N}$ dei numeri naturali. Il sistema comunemente accettato è quello dovuto a *Peano* ed assiomatizza l'idea intuitiva che l'insieme dei numeri naturali è la conclusione del processo di "aggiungere un nuovo elemento (il successore) ad un elemento dato". Per Peano i numeri naturali sono dunque un insieme $\mathbf{N}$, con un elemento privilegiato 0 (tradizionalmente chiamato "zero") ed una funzione

$$\sigma\colon \mathbf{N} \longrightarrow \mathbf{N}$$

(detta "successore"), che soddisfano gli assiomi:

- P_1) 0 *non è successore di alcun naturale*: $0 \neq \sigma(n)$, per ogni $n \in \mathbf{N}$;
- P_2) *la funzione successore è iniettiva*: se $\sigma(n) = \sigma(m)$, allora $n = m$;
- P_3) *principio di induzione*: per dimostrare che un sottoinsieme $U \subseteq \mathbf{N}$ coincide con tutto $\mathbf{N}$, basta che si verifichino i seguenti due fatti:

i) $0 \in U$;

ii) se $n \in U$, allora anche $\sigma(n) \in U$, per ogni $n \in \mathbf{N}$.

Vedremo nel seguito come solo a partire da questi assiomi si possono dedurre tutte le principali proprietà dei numeri naturali. Per ora, osserviamo solo che i primi due assiomi ci dicono che l'insieme $\mathbf{N}$ non è finito, dunque che è *infinito*, poichè la funzione σ è una endofunzione che è iniettiva, ma non suriettiva.

Una prima importante conseguenza degli assiomi di Peano è la possibilità di *definire* funzioni $\mathbf{N} \longrightarrow X$ (successioni) mediante un procedimento effettivo a partire da dati iniziali, nel modo seguente:

Teorema 2.1.1 *("definizione per induzione"). Se X è un qualsiasi insieme, $a \in X$ è un elemento scelto di X e $f: X \longrightarrow X$ è un qualsiasi endomorfismo di X, allora esiste un'unica funzione $h: \mathbf{N} \longrightarrow X$ tale che:*

1) $h(0) = a$;

2) $h(\sigma(n)) = f(h(n))$.

Non diamo la dimostrazione di questo teorema. Pensando alla funzione σ come alla funzione $\sigma(n) = n + 1$, la funzione $h: \mathbf{N} \longrightarrow X$ la cui esistenza e unicità è garantita dal teorema è la funzione $h(n) = f^n(a)$; dunque il fatto che questo teorema sia una conseguenza degli assiomi di Peano è abbastanza chiaro, anche se una dimostrazione formale dell'esistenza della funzione h è un po' complicata; la sua unicità è invece immediata: se esistesse un'altra funzione $h': \mathbf{N} \longrightarrow X$ tale $h'(0) = a$ e $h'(\sigma(n)) = f(h(n))$, allora sia $U = \{n \in \mathbf{N} | h(n) = h'(n)\}$; si ha: $0 \in U$, poichè $h'(0) = a = h(0)$;inoltre, se $n \in U$, allora anche $\sigma(n) \in U$, perchè $h(\sigma(n)) = \sigma(h(n)) = \sigma(h'(n)) = h'(\sigma(n))$; dunque per il principio di induzione possiamo concludere che $U = \mathbf{N}$, quindi che $h(n) = h'(n)$ per ogni $n \in \mathbf{N}$; perciò $h = h'$.

L'uso più o meno esplicito della definizione di funzioni per induzione è assai frequente in matematica. Ad esempio, se $(M, \cdot, e)$ è un monoide e se $a \in M$, le "*potenze n-esime*" di a sono in realtà una funzione $a^{(\ )}: \mathbf{N} \longrightarrow M$ (la cui esistenza ed unicità è garantita dal precedente teorema nel modo seguente: si consideri la funzione $f: M \longrightarrow M$ data

dal prodotto per a, cioè $f(x) = xa$; applicando il teorema si ha che esiste una ed una sola funzione $h = a^{(\ )}\colon \mathbf{N} \longrightarrow \mathbf{N}$ che soddisfa

$$a^0 = e\,, \quad a^{\sigma(n)} = a^n a\,.$$

Il teorema sulla definizione per induzione è un enunciato di universalità per il diagramma

$$* \xrightarrow{\ 0\ } \mathbf{N} \xrightarrow{\ \sigma\ } \mathbf{N}\,:$$

per *ogni* altro diagramma della stessa forma:

$$* \xrightarrow{\ x\ } X \xrightarrow{\ f\ } X$$

esiste un'unica funzione (successione) $h\colon \mathbf{N} \longrightarrow X$ tale che i seguenti diagrammi commutino:

$$\begin{array}{ccccc} * & \xrightarrow{\ 0\ } & \mathbf{N} & \xrightarrow{\ \sigma\ } & \mathbf{N} \\ \| & & \downarrow{\scriptstyle h} & & \downarrow{\scriptstyle h} \\ * & \xrightarrow{\ x\ } & X & \xrightarrow{\ f\ } & X\,. \end{array}$$

Si può dimostrare di più: il teorema sulla definizione per induzione è in realtà *equivalente* agli assiomi di Peano; dunque questa proprietà universale del diagramma $* \xrightarrow{\ 0\ } \mathbf{N} \xrightarrow{\ \sigma\ } \mathbf{N}$ costituisce un'altra formulazione più semplice degli assiomi per i numeri naturali ("*assioma di Peano-Lawvere*").

Si osservi che dal teorema di definizione per induzione si ha immediatamente la *unicità* dell'insieme dei numeri naturali. Più precisamente, se $(0', \mathbf{N}', \sigma')$ è un altro insieme munito di un elemento $0'$ e di un endomorfismo σ' che soddisfano gli assiomi di Peano, allora esiste un unico *isomorfismo* $h\colon \mathbf{N} \longrightarrow \mathbf{N}'$ tale che $h(0) = 0'$ e $\sigma' h = h\sigma$. Infatti, applicando il teorema sulla definizione per induzione a $(0', \mathbf{N}', \sigma')$ si ha che esiste un'unica funzione $h\colon \mathbf{N} \longrightarrow \mathbf{N}'$ tale che $h(0) = 0'$ e $h(\sigma(n)) = \sigma'(h(n))$; poichè anche $\mathbf{N}'$ soddisfa gli assiomi di Peano, anche per $\mathbf{N}'$ si può dimostrare il teorema di definizione per induzione;

applicando tale teorema a $(0, \mathbf{N}, \sigma)$, si ha che esiste una sola funzione $h' \colon \mathbf{N}' \longrightarrow \mathbf{N}$ tale che $h'(0') = 0$ e $\sigma(h'(n)) = h'(\sigma'(n))$. Consideriamo la composizione $k = h'h \colon \mathbf{N} \longrightarrow \mathbf{N}$; poichè $k(0) = h'(h(0)) = h'(0') = 0$ e $k(\sigma(n)) = h'(h(\sigma(n)) = h'(\sigma'(h(n))) = \sigma(h'(h((n))) = \sigma(k(n))$ e poichè anche l'identità $1_{\mathbf{N}} \colon \mathbf{N} \longrightarrow \mathbf{N}$ soddisfa le stesse condizioni, l'unicità espresssa dal teorema di definizione per induzione assicura che $k = h'h = 1_{\mathbf{N}}$. Un ragionamento analogo prova che anche la composizione hh' è l'identità di $\mathbf{N}'$; dunque h è un isomorfismo, essendo h' la sua funzione inversa.

2.2 Ricorsività

La classe delle funzioni

$$\mathbf{N} \times \mathbf{N} \times \ldots \times \mathbf{N} \longrightarrow \mathbf{N}$$

che si possono definire per induzione e per composizione di funzioni a partire da 0, dalla funzione successore $\sigma \colon \mathbf{N} \longrightarrow \mathbf{N}$ e dalle proiezioni è detta classe delle *funzioni ricorsive primitive*. Si osservi come la clausola riguardante la chiusura rispetto alla composizione ci permette di ottenere i numeri naturali come $0, 1 = \sigma(0), 2 = \sigma(\sigma(0)), \ldots$.

Vediamo ora come tutte le funzioni ordinarie della aritmetica appartengano alla classe delle funzioni ricorsive primitive e come le loro principali proprietà possano essere dimostrate solo a partire dagli assiomi di Peano.

La possibilità di definire le potenze n-esime di un elemento di un monoide applicata all'elemento σ del monoide $\mathrm{End}(\mathbf{N})$ in cui, ricordiamo, l'operazione è la composizione, fornisce la *somma*; definiamo dunque la somma $m + n$, per ogni m fissato, come:

$$m + n = \sigma^n(m) \, .$$

Ricordando la definizione induttiva delle potenze, questa definizione equivale alla definizione induttiva:

i) $m + 0 = \sigma^0(m) = 1_{\mathbf{N}}(m) = m$;

ii) $m + \sigma(n) = (\sigma^n(m)) = \sigma(m + n)$.

Similmente, definiamo il *prodotto* mn, per ogni m fissato, come

$$mn = (\sigma^m)^n(0)\,.$$

Anche questa definizione equivale alla definizione induttiva:

i) $m0 = 0$;

ii) $m\sigma(n) = (\sigma^m)^{\sigma(n)}(0) = [\sigma^m\sigma^{mn}](0) = \sigma^m(mn) = mn + m$.

La definizione di somma data sopra implica che $\sigma(n) = n+1$. È dunque chiaro come queste definizioni coincidano nel modello intuitivo dei numeri naturali alle abituali operazioni di somma e prodotto. Ma per poter *dimostrare* che queste definizioni danno delle operazioni che hanno le usuali proprietà della somma e del prodotto di naturali, bisogna prima dimostrare alcune elementari proprietà delle potenze degli elementi di un monoide.

Lemma 2.2.1 *Siano a, b due elementi di un monoide $(M, \cdot, e)$. Se a e b sono permutabili ($ab = ba$), allora:*

1) *per ogni coppia di naturali n, m si ha: $a^n b^m = b^m a^n$;*

2) *per ogni naturale n si ha: $(ab)^n = a^n b^n$.*

Dimostrazione. 1) Dimostriamo dapprima che per ogni $n \in \mathbf{N}$ si ha $a^n b = ba^n$. Questo equivale a dimostrare che l'insieme:

$$U = \{n \in \mathbf{N} \mid a^n b = ba^n\}$$

coincide con l'insieme di *tutti* i numeri naturali. Applichiamo dunque il principio di induzione: $0 \in U$, poichè $a^0 = e$; se $n \in U$, cioè se $a^n b = ba^n$, allora $a^{\sigma(n)}b = a^n ab = a^n ba = ba^n a = ba^{\sigma(n)}$, dunque anche $\sigma(n) \in U$; perciò $U = \mathbf{N}$. Per dimostrare che $a^n b^m = b^m a^n$ per tutti i naturali m e n, basta dimostrare che, posto $c = a^n$, si ha $cb^m = b^m c$, per tutti gli $m \in \mathbf{N}$; per quanto già dimostrato, questo è vero se $cb = bc$, cioè se $a^n b = ba^n$; per quanto già dimostrato, questo è vero perchè $ab = ba$.

2) Procediamo per induzione. Per $n = 0$ la proprietà è vera perchè $a^0 = b^0 = e$; supponiamo che la proprietà sia vera per un certo n e

dimostriamola per il successivo $\sigma(n)$: $(ab)^{\sigma(n)} = (ab)(ab)^n = aba^n b^n = aa^n bb^n = a^{\sigma(n)} b^{\sigma(n)}$. ■

Si osservi che in particolare ogni elemento a è permutabile con se stesso; dunque per ogni n e m si ha: $a^n a^m = a^m a^n$.

Lemma 2.2.2 . *Per ogni coppia di naturali n e m e per ogni elemento a di un monoide si ha: $(a^n)^m = (a^m)^n$.*

DIMOSTRAZIONE. Sia

$$U = \{n \in \mathbf{N} \mid (a^n)^m = (a^m)^n, \text{per tutti gli } m \in \mathbf{N}\}.$$

Si ha $0 \in U$, perchè per tutti gli $m \in \mathbf{N}$ vale $(a^0)^m = e^m = e = (a^m)^0$. Se $m \in U$,cioè se n è tale che per tutti gli $m \in \mathbf{N}$ si ha $(a^n)^m = (a^m)^n$, allora:
$(a^m)^{\sigma(n)} = (a^m)(a^m)^n = a^m(a^n)^m = (aa^n)^m =$
(perchè a e a^n sono permutabili) $= (a\sigma(n))^m$
per ogni $m \in \mathbf{N}$; dunque $\sigma(n) \in U$, perciò $U = \mathbf{N}$. ■

Si presti attenzione al senso di ciò che si sta facendo. Può sembrare ovvio che $(a^n)^m = (a^m)^n$, perchè l'esperienza che abbiamo dei numeri naturali ci dice che entrambe tali quantità sono uguali ad a^{nm}. Ma procedendo da una definizione assiomatica dei numeri naturali, a questo punto non sappiamo ancora che cosa sono la somma ed il prodotto di numeri naturali. Tali operazioni non compaiono nella definizione dei numeri naturali che abbiamo dato, perciò dobbiamo dedurre la loro esistenza (ciò che abbiamo fatto) e le loro proprietà che le qualificano come "la somma ed il prodotto" (ciò che non abbiamo ancora fatto). Anzi, useremo le proprietà delle potenze appena dimostrate proprio per dimostrare che le definizioni di somma e prodotto di numeri naturali che abbiamo dato soddisfano le proprietà che ci aspettiamo debbano soddisfare.

Teorema 2.2.1 . *L'insieme $\mathbf{N}$ dei numeri naturali è un monoide commutativo per le operazioni di somma e prodotto precedentemente definite, per cui valgono le seguenti proprietà:*

1) $(\mathbf{N}, +, 0)$ *e* $(\mathbf{N}, \cdot, 1)$ *hanno la proprietà di cancellazione: se $n+p = m+p$, allora $n = m$ e, se $p \neq 0$ e $np = mp$, allora $n = m$; inoltre, se $m + n = 0$, allora $m = 0$ e $n = 0$;*

2) *vale l'identità:* $m(n+p) = mn + mp$ *(proprietà distributiva);*

3) $mn = 0$ *se e solo se* $m = 0$ *o* $n = 0$*;*

4) *per ogni monoide* $(M, \cdot, e)$ *e per ogni elemento* $a \in M$*, si ha:*

$$a^n \cdot a^m = a^{n+m}, \quad (a^m)^n = a^{mn}.$$

DIMOSTRAZIONE. A titolo di esempio dimostriamo solo alcune delle proprietà enunciate. Dimostriamo che 0 è l'elemento neutro per la somma, dunque che per ogni $n \in \mathbf{N}$ si ha $n + 0 = n$. Per induzione su n: la proprietà è vera per $n = 0$, perchè $\sigma^0(0) = 1_{\mathbf{N}}(0) = 0$; supponiamo che la proprietà sia vera per n e dimostriamola per $\sigma(n)$:
$\sigma^{\sigma(n)}(0) = \sigma(\sigma^n(0)) = \sigma(n)$.

Dimostriamo la commutatività della somma, usando il primo dei due lemmi precedenti:
$n + m = \sigma^n(m) = \sigma^n(\sigma^m(0)) = (\sigma^n\sigma^m)(0) = (\sigma^m\sigma^n)(0) = \sigma^m(n) = m + n$.

Dimostriamo la proprietà di cancellazione per la somma: se $m+n = p + n$, allora $m = p$. Dimostriamo per induzione su n, per tutti gli m e p; se $n = 0$, la proprietà è vera; supponiamo che la proprietà sia vera per un certo n e per tutti gli m ed i p, e dimostriamola per $\sigma(n)$: se $m + \sigma(n) = p + \sigma(n)$, allora $\sigma(m + n) = \sigma(p + n)$; poichè σ è iniettiva, si ha $m + n = p + n$; per l'ipotesi di induzione si ha $m = p$.

Dimostriamo la prima delle proprietà 4). Dimostriamo per induzione su n che $a^{n+m} = a^n \cdot a^m$, per ogni $m \in \mathbf{N}$. La proprietà è vera per $n = 0$; supponiamo che sia vera per un certo n e dimostriamola per il successivo $\sigma(n)$:
$a^{\sigma(n)} \cdot a^m = a^n \cdot a \cdot a^m = a^n \cdot a^m \cdot a = a^{n+m} \cdot a = a^{\sigma(n+m)} = a^{\sigma(n)+m}$.

In modo simile, sfruttando gli assiomi o i lemmi già dimostrati, si dimostrano tutte le rimanenti proprietà. ■

A partire dalla somma e dal prodotto si possono definire per induzione e composizione molte altre funzioni. Un esempio è la seguente funzione (detta "successione di Fibonacci"):
$f(0) = 1, f(1) = 1, f(\sigma(\sigma(n)) = f(\sigma(n)) + f(n)$,
cioè, interpretando $\sigma(n)$ come $n + 1$,
$f(0) = 1, f(1) = 1, f(n + 2) = f(n + 1) + f(n)$.

I primi termini di questa successione sono: $1, 1, 2, 3, 5, 8, 13, \ldots$. È chiaro che tutte le funzioni definite in questo modo, che abbiamo convenuto di chiamare funzioni ricorsive primitive, hanno il requisito di essere effettivamente calcolabili, cioè calcolabili con una procedura effettiva che può in principio essere eseguita da qualsiasi macchina. Tuttavia si può dimostrare che esistono altre funzioni che hanno questo requisito, ma che non sono ricorsive primitive. Ad esempio la funzione

$$f\colon \mathbf{N} \times \mathbf{N} \longrightarrow \mathbf{N}$$

definita secondo lo schema seguente:
$f(0, m) = \sigma(m); f(n, 0) = \sigma(n); f(\sigma(n), \sigma(m)) = f(n, f(\sigma(n), m))$,
è chiaramente una funzione che può essere calcolata da una qualsiasi macchina (se ne calcolino alcuni valori o, meglio, si scriva un programma che la calcola). Tuttavia si può dimostrare che non appartiene alla classe delle funzioni ricorsive primitive, ma ad una classe piu ampia chiamata classe delle "*funzioni ricorsive (generali)*". Tale classe è suscettibile di una definizione precisa ed il suo studio è il contenuto della teoria della ricorsività. La "*tesi di Church*" afferma che la classe delle funzioni ricorsive coincide con la classe delle funzioni effettivamente calcolabili. Tale tesi non è dimostrabile poichè da una parte si ha un concetto formale, come quello di funzione ricorsiva, mentre dall'altra si ha un concetto non formale, come quello di "funzione effettivamente calcolabile". Tuttavia a tutt'oggi non sono state trovate funzioni calcolabili con una procedura effettiva che non siano anche ricorsive.

2.3 Ordine e Divisione

Siamo ora in grado di descrivere la ulteriore struttura di cui $\mathbf{N}$ è dotato e di provarne le proprietà, sempre deducendole dagli assiomi di Peano o dalle loro conseguenze che abbiamo già dimostrato.

La "*relazione d'ordine*" di $\mathbf{N}$ si definisce nel modo seguente:

$$n \leq m \text{ se e solo se esiste } k \text{ tale che } n + k = m\,;$$

Scriveremo $n < m$ per $n \leq m$ e $n \neq m$, dunque se esiste $k \neq 0$ tale che $n + k = m$. È immediato dimostrare che la relazione $n \leq m$ è una

relazione d'ordine (si veda il paragrafo 6 del capitolo 3 per la definizione precisa di relazione d'ordine). Le proprietà particolari della relazione d'ordine su $\mathbf{N}$ sono espresse dal

Teorema 2.3.1 i) *legge di "tricotomia": per ogni $m, n \in \mathbf{N}$ vale una sola delle seguenti possibilità: $m < n, m = n, n < m$.*

ii) *se $m \leq n$, allora $m + k \leq n + k$ e $mk \leq nk$, per ogni k.*

iii) $\mathbf{N}$ *è "bene ordinato": ogni sottoinsieme $V \subseteq \mathbf{N}$ non vuoto ha un primo elemento, cioè un elemento $v \in V$ tale che $v \leq x$, per ogni elemento $x \in V$.*

DIMOSTRAZIONE. Dimostriamo iii) a titolo di esempio. Procediamo per assurdo; supponiamo che $V \subseteq \mathbf{N}$ sia un sottoinsieme non vuoto che non abbia un primo elemento e dimostriamo che in tal caso ogni naturale è minore o uguale ad ogni elemento di V, cioè che l'insieme

$$U = \{n \,|\, x \in V \Rightarrow n \leq x\}$$

coincide con l'insieme di tutti i naturali. Procediamo per induzione; $0 \in U$; se $n \in U$, allora $n \notin V$, perchè in caso contrario V avrebbe n come primo elemento; dunque se $x \in V$, allora $n \leq x$ e $x \neq n$; ma se $n \leq x$ e $x \neq n$, allora $\sigma(n) \leq x$, dunque anche $\sigma(n) \in U$. Dunque $U = \mathbf{N}$. Ma questa è una contraddizione, perchè se V non è vuoto, sia $k \in V$; allora $n \leq k$, per ogni n, dunque anche per $n = k + 1$, assurdo. ■

La prima importante applicazione dell'ordine sui naturali è il teorema sulla divisione di naturali.

Teorema 2.3.2 . *Per ogni coppia di naturali x e n, $n \neq 0$, esistono due naturali q ("quoziente") e r ("resto"), tali che:*

$$x = qn + r \quad e \quad 0 \leq r < n\,.$$

Inoltre, se q' e r' sono altri due naturali per cui $0 \leq r' < n$ e $x = q'n + r'$, allora $q = q'$ e $r = r'$ (unicità del quoziente e del resto della divisione per n).

DIMOSTRAZIONE. Consideriamo l'insieme dei "possibili resti":

$$M = \{m \,|\, \text{esiste } q \text{ tale che } x = qn + m\}\,.$$

Definiamo r come il primo elemento di M; bisogna però assicurarsi che M non sia vuoto; infatti, x stesso sta in M (basta prendere $q = 0$). Dunque esiste q tale che $x = qn + r$. Dobbiamo dimostrare che $0 \leq r < n$. Per assurdo, supponiamo $n \leq r$; allora $r = n + k$, quindi $x = qn + n + k = (q + 1)n + k$ e $k < r$; assurdo perchè r è il primo elemento di M.

Se $x = q'n + r'$, con $0 \leq r' < n$, dimostriamo che r' deve essere un primo elemento di M. Infatti, se così non fosse, esisterebbe $m \in M$ con $m < r'$; dunque $r' = m + y$ ed esisterebbe q tale che $x = qn + m = q'n + r' = q'n + m + y$; quindi $qn = q'n + y$, perciò y sarebbe un multiplo di n, ciò che è impossibile perchè $r' < n$. Dunque $r = r'$, perciò per la cancellazione della somma anche $qn = q'n$ e, per la cancellazione del prodotto, poichè $n \neq 0$, anche $q = q'$. ■

2.4 Idempotenti ed involuzioni

Per ogni endofunzione $f\colon X \longrightarrow X$ su un insieme X possiamo definire il sottoinsieme di X

$$F = Fix(f) = \{x \in X | f(x) = x\} \subseteq X$$

dei *punti fissi* di f. Se f è un *idempotente*, cioè ha la proprietá che $ff = f^2 = f$, dunque che $f(f(x)) = f(x)$ per ogni $x \in X$, allora ogni elemento della forma $f(x)$ è un punto fisso, perchè $y = f(x)$ ha la proprietà che $f(y) = y$. Dunque denotando con e la funzione f stessa considerata però come funzione $X \longrightarrow F$ e denotando con $i\colon F \hookrightarrow X$ l'inclusione di F in X considerata come una particolare funzione (la restrizione dell'identità su X), si ha che l'idempotente f si fattorizza come

$$X \xrightarrow{\ f\ } X \qquad = \qquad X \xrightarrow{\ e\ } F \xhookrightarrow{\ i\ } X\,.$$

Cosa possiamo dire della composizione

$$F \xhookrightarrow{\ i\ } X \xrightarrow{\ e\ } F\ ?$$

È chiaro che

$$F \overset{i}{\hookrightarrow} X \xrightarrow{e} F \quad = \quad F \xrightarrow{1_F} F\,,$$

poichè se $y \in F$, cioè se $y = f(y)$, allora $e(i(y)) = e(y) = f(y) = y$. Diremo che la coppia e, i è uno *spezzamento* dell'idempotente f. Diremo anche che $i\colon F \hookrightarrow X$ è un *retratto* e che e è una sua *retrazione* quando la composizione ei è l'identità su F. Dunque ogni idempotente determina un retratto, prendendone lo spezzamento attraverso i punti fissi. Viceversa, dato un retratto $i\colon F \hookrightarrow X$, *ogni* sua retrazione $e\colon X \longrightarrow F$ determina un idempotente $f = ie\colon X \longrightarrow X$ (poichè $f^2 = (ie)(ie) = i(ei)e = ie = f$) che ha F come insieme dei punti fissi. Da questa discussione dovrebbe apparire chiaro che se $i\colon F \hookrightarrow X$, allora le sue retrazioni $e\colon X \longrightarrow F$, quindi gli idempotenti che hanno F come insieme dei punti fissi, sono in corrispondenza biunivoca con le funzioni $F^c \longrightarrow F$ dall'insieme complementare dell'insieme F dei punti fissi a F stesso. In particolare, se X è un insieme finito di cardinalità n, allora gli idempotenti che hanno un insieme di punti fissi di cardinalità i sono

$$\binom{n}{i} i^{n-i}$$

e quindi il numero di *tutti* gli idempotenti su X è

$$\sum_{i=0}^{n} \binom{n}{i} i^{n-i}\,.$$

Un'altra importante proprietà che un endomorfismo $f\colon X \longrightarrow X$ può avere è quella di essere una *involuzione*, cioè $f^2 = 1$ e dunque $f(f(x)) = x$, per ogni $x \in X$. Si osservi che mentre l'unico idempotente che sia anche invertibile è l'identità, invece *ogni* involuzione è invertibile, essendo essa stessa il proprio inverso. Per capire meglio che cosa sia una involuzione cercheremo di contare il numero delle involuzioni I_n su un insieme finito di cardinalità n. Una formula induttiva per il numero delle involuzioni su un insieme finito di cardinalità n è

$$I_n = I_{n-1} + (n-1)I_{n-2}$$

e $I_0 = 1 = I_1$. Infatti, se $n \geq 2$, ripartiamo l'insieme delle involuzioni $f\colon [n] \longrightarrow [n]$ in due classi disgiunte, quelle che lasciano fisso 1 e quelle

che lo muovono. Il numero delle prime è chiaramente I_{n-1}, perchè una tale involuzione consiste semplicemente nel dare una involuzione sui rimanenti $n-1$ elementi. Per determinare il numero delle seconde, si osservi che se $f(1) \neq 1$, allora f manda la coppia $\langle 1, f(1)\rangle$ nella coppia $\langle f(1), 1\rangle$, *perchè f è una involuzione*; dunque f è completamente determinata da una involuzione sui rimanenti $n-2$ elementi. Poichè $f(1)$ può assumere solo $n-1$ valori, per la clausola $f(1) \neq 1$, il numero delle seconde è $(n-1)I_{n-2}$.

2.5 Funzioni generatrici

Usiamo il problema di trovare una espressione esatta per il numero I_n delle involuzioni su un insieme con n elementi, per illustrare un metodo classico della combinatoria, il *metodo delle funzioni generatrici*, che consiste nel trovare una serie formale (di Taylor)

$$I(t) = \sum_{n\geq 0} I_n \frac{t^n}{n!}$$

che abbia per coefficienti proprio la successione dei numeri I_n, usando la formula induttiva appena stabilita. Nel caso in questione, per la forma della formula induttiva, tentiamo di trovare la *derivata* $\dot{I}(t)$ di $I(t)$, supponendo che una tale funzione esista. Si ha:

$\dot{I}(t) = \sum_{n\geq 1} I_n \frac{t^{n-1}}{(n-1)!} = 1 + \sum_{n\geq 2} I_n \frac{t^{n-1}}{(n-1)!} =$

$= 1 + \sum_{n\geq 2} I_{n-1} \frac{t^{n-1}}{(n-1)!} + \sum_{n\geq 2}(n-1) I_{n-2} \frac{t^{n-1}}{(n-1)!} =$

$= 1 + \sum_{n\geq 2} I_{n-1} \frac{t^{n-1}}{(n-1)!} + t \sum_{n\geq 2} I_{n-2} \frac{t^{n-2}}{(n-2)!} =$

$= 1 + [I(t) - 1] + tI(t) = I(t) + tI(t) = I(t)(1+t).$

Poichè $\frac{\dot{I}(t)}{I(t)} = 1 + t$ si ha $\log I(t) = t + \frac{t^2}{2}$; dunque la funzione generatrice del numero delle involuzioni è

$$I(t) = e^{t+\frac{t^2}{2}} .$$

Non rimane dunque che sviluppare in serie di Taylor questa funzione per trovare i valori esatti di I_n. Osserviamo dapprima che la proprietà fondamentale dell'esponenziale è che *trasforma somme in prodotti*, dunque

che $e^{t+\frac{t^2}{2}} = e^t\, e^{\frac{t^2}{2}}$; ricordiamo anche che $e^x = \sum_{i\geq 0} \frac{x^i}{i!}$ e che il *prodotto di Cauchy* $\sum_{i\geq 0} a_i t^i \sum_{k\geq 0} b_k t^k$ di due serie è la serie $\sum_{n\geq 0} c_n t^n$, dove il coefficiente c_n è dato dalla formula $c_n = \sum_{i+k=n} a_i b_k$, ciò che generalizza il prodotto di polinomi. Si ha dunque che

$\frac{I_n}{n!} t^n = \sum_{i\geq 0} \frac{t^i}{i!} \sum_{k\geq 0} \frac{(\frac{t^2}{2})^k}{k!} = \sum_{i\geq 0} \frac{t^i}{i!} \sum_{k\geq 0} \frac{t^{2k}}{k!2^k} =$

$= \sum_{n\geq 0} (\sum_{i+2k=n} \frac{1}{i!k!2^k})\, t^n$

e quindi che

$$I_n = \sum_{i+2k=n} \frac{n!}{i!k!2^k}\,.$$

2.6 Esercizi

1. I *numeri di Fibonacci* F_n sono definiti induttivamente da $F_0 = F_1 = 1$ e $F_n = F_{n-1} + F_{n-2}$, per $n \geq 2$, ed hanno svariate interpretazioni combinatorie. Si provi che la funzione generatrice $F(t)$ dei numeri di Fibonacci è

   $$F(t) = \frac{1}{1 - t - t^2}\,.$$

   (Suggerimento: si assuma che $F(t)$ sia sviluppabile in serie di potenze, $F(t) = \sum_{n\geq 0} F_n t^n$; allora

   $F(t) = 1 + t + \sum_{n\geq 2} F_n t^n = 1 + t + \sum_{n\geq 2} F_{n-1} t^n + \sum_{n\geq 2} F_{n-2} t^n =$

   .....)

2. Ricordando che la funzione $\frac{1}{1-x}$ è sviluppabile in serie tramite la *serie geometrica*

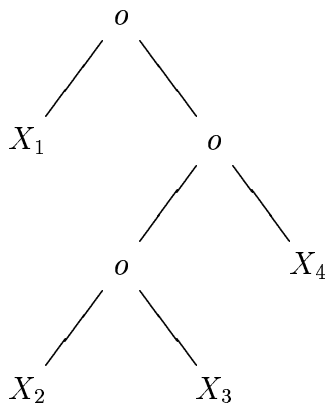
   $$\frac{1}{1-x} = \sum_{k\geq 0} x^k\,,$$

   si usi il precedente esercizio per mostrare che i numeri di Fibonacci sono dati dalla formula

   $$F_n = \sum_{i+k=n} \binom{k}{i}\,.$$

   (Suggerimento: sostituire $t(1+t)$ ad x nella serie geometrica ed usare la formula del binomio di Newton).

3. Ricordiamo che nell'esercizio 1.13.6 abbiamo introdotto i numeri di Catalano a_n e che abbiamo proposto di provare che la loro definizione induttiva è $a_n = \sum_{k=1}^{n-1} a_k a_{n-k}$, per $n \geq 2$, convenendo che $a_1 = 1 = a_2$; conveniamo anche che $a_0 = 0$ (trovare una giustificazione per tali convenzioni, anche sulla base di quanto segue). Definiamo *albero binario* un grafo con la proprietà che da ogni nodo partono due rami o nessuno; ad esempio



è un albero binario con 4 *foglie* (= nodi terminali). Numerando le foglie in senso antiorario come nella figura e convenendo che un nodo non terminale è il prodotto dei due nodi immediatamente sottostanti, si puo' associare ad ogni tale grafo una configurazione di parentesi; ad esempio, quella associata al grafo in figura è $X_1((X_2X_3)X_4)$. Si dimostri che il numero di Catalano a_n è il numero degli alberi binari con n foglie, stabilendo una opportuna corrispondenza biunivoca.

4. Si dimostri che la funzione $A(t)$ generatrice dei numeri di Catalano è

$$A(t) = \frac{1}{2}(1 - \sqrt{1-4t}) \ .$$

(Suggerimento: $A(t) = \sum_{n\geq 0} a_n t^n = t + \sum_{n\geq 2} a_n t^n =$

$= t + [\sum_{n\geq 2}(\sum_{i=1}^{n-1} a_{n-i}a_i)t^n] = t + [\sum_{n\geq 2}(\sum_{i+k=n} a_i a_k)t^n] = t + A^2(t)$, quindi... Infine si tenga conto della condizione iniziale $A(0) = a_0 = 0$.)

5. Si trovi lo sviluppo in serie di Taylor della funzione generatrice dei numeri di Catalano. (Suggerimento: si dimostri per induzione che per ogni $n \geq 2$, la derivata n-esima di $A(t)$ è

$$A^{(n)}(t) = \frac{1}{2^n}[\prod_{k=0}^{n-2}(2k+1)](\frac{1}{4} - t)^{-\frac{2n-1}{2}} \quad).$$

Usando poi che $n!a_n = A^{(n)}(0)$ si deduca l'espressione esatta

$$a_n = \frac{1}{n}\binom{2n-2}{n-1},$$

moltiplicando numeratore e denominatore di

$$a_n = \frac{2^{2n-1}}{n!2^n}\prod_{k=0}^{n-2}(2k+1)$$

per la quantità $\prod_{k=1}^{n-1}(2k) = 2^{n-1}(n-1)!)$.

6. Se

$$X \xrightarrow{f} X \quad = \quad X \xrightarrow{e} F \xhookrightarrow{i} X \quad =$$

$$= \quad X \xrightarrow{p} G \xhookrightarrow{j} X$$

sono due spezzamenti dello stesso idempotente f su X, si provi che esiste un unico isomorfismo $t\colon F \longrightarrow G$ tale che $jt = i$ (o equivalentemente, $te = p$).

2.7 I numeri interi

Vogliamo ora mostrare come le proprietà dell'insieme **N** dei numeri naturali permettano di costruire l'insieme **Z** degli interi, come soluzione al problema di "aggiungere i numeri *negativi*", cioè gli opposti di ogni numero naturale. Cercheremo anche di spiegare perchè il prodotto di due negativi è positivo ("meno per meno fa più"). Ricordiamo che l'ordine stretto su **N** è definito da $n < m$ se esiste un $d \neq 0$ tale che

$n + d = m$; ricordiamo anche che un tale d è *unico*, per la proprietà di cancellazione della somma, ed è da pensare come "$m - n$" o, meglio, come l'unica soluzione dell'equazione $n + x = m$. Consideriamo ora il prodotto $\mathbf{N} \times \mathbf{N}$ e la funzione

$$\mathbf{N} \times \mathbf{N} \xrightarrow{f} \mathbf{N} \times \mathbf{N}$$

definita da

$$f(n_1, n_2) = \begin{cases} \langle 0, d \rangle & \text{se } n_1 < n_2 \text{ e se } n_1 + d = n_2 \\ \langle d, 0 \rangle & \text{se } n_2 < n_1 \text{ e se } n_2 + d = n_1 \\ \langle 0, 0 \rangle & \text{se } n_1 = n_2 \end{cases}$$

Tale funzione è ben definita per la proprietà di *tricotomia* dell'ordine in $\mathbf{N}$. È facile inoltre verificare che f è *idempotente*. Sia

$$\mathbf{N} \times \mathbf{N} \xrightarrow{e} \mathbf{Z} \xhookrightarrow{i} \mathbf{N} \times \mathbf{N}$$

lo spezzamento di f attraverso l'insieme $\mathbf{Z}$ dei suoi *punti fissi*. Dunque

$$\mathbf{Z} = \{\langle n_1, n_2 \rangle \mid n_1 = 0 \quad \text{o} \quad n_2 = 0\} .$$

Possiamo convenire di denotare con $-n$ le coppie del tipo $\langle n, 0 \rangle$, con $n \neq 0$ e semplicemente con n le coppie del tipo $\langle 0, n \rangle$, con $n \in \mathbf{N}$. Dunque come insieme, $\mathbf{Z}$ così definito coincide con l'insieme dei numeri interi ("relativi") che conosciamo da sempre. Il punto è di mostrare come questa descrizione implica l'esistenza delle usuali operazioni sugli interi e delle loro usuali proprietà. Osserviamo dapprima che le operazioni di somma e prodotto su $\mathbf{N}$ si possono estendere alle coppie ordinate $\mathbf{N} \times \mathbf{N}$ definendole "*puntualmente*", cioè
$\langle n_1, n_2 \rangle + \langle m_1, m_2 \rangle = \langle n_1 + m_1, n_2 + m_2 \rangle$
$\langle n_1, n_2 \rangle \langle m_1, m_2 \rangle = \langle n_1 m_1, n_2 m_2 \rangle$.
È facile verificare che in tal modo si ottengono due operazioni sul prodotto $\mathbf{N} \times \mathbf{N}$ che soddisfano le stesse proprietà formali delle analoghe operazioni su $\mathbf{N}$, cioè *associatività*, esistenza dell'*elemento neutro*, *commutatività*. Ciò si esprime dicendo che $\mathbf{N} \times \mathbf{N}$ ha, come $\mathbf{N}$, due strutture di *monoide commutativo*. Inoltre, le due operazioni di somma e prodotto soddisfano la stessa proprietà di *distributività* che vale in $\mathbf{N}$:

$$x(y+z) = xy + xz$$

$$x0 = 0 \, .$$

Esprimiamo tutto ciò dicendo che $\mathbf{N}\times\mathbf{N}$ ha una struttura di *semi-anello commutativo*, definita puntualmente dalla struttura di semi-anello commutativo di $\mathbf{N}$. Vedremo che tale struttura di semi-anello commutativo *non* è l'unica possibile su $\mathbf{N} \times \mathbf{N}$, ma ne esiste un'altra che viene "ereditata" da $\mathbf{Z}$ (si veda l'esercizio 2.13.1). Il fatto generale che governa questi fenomeni è descritto nel seguente

Teorema 2.7.1 *Sia $(X, *, u)$ un monoide e sia*

$$f\colon X \longrightarrow X$$

una funzione idempotente. *Sia*

$$X \xrightarrow{\quad f \quad} X \qquad = \qquad X \xrightarrow{\quad e \quad} F \xhookrightarrow{\quad i \quad} X$$

lo spezzamento di f attraverso i punti fissi. Allora F ha una unica struttura di monoide $(F, \circ, u')$ per cui la proiezione $e\colon X \longrightarrow F$ è un omomorfismo, *se e solo se f soddisfa la seguente identità (che chiameremo proprietá di* chiusura *rispetto all'operazione $*$):*

$$f[f(x) * f(y)] = f(x * y) \, .$$

Inoltre, tutte le identità che sono soddisfatte dal monoide $(X, *, u)$ sono anche soddisfatte dal monoide $(F, \circ, u')$ e, se X ha altre strutture di monoide per cui f soddisfa la proprietà di chiusura, allora tutte le identità che sussistono tra le varie operazioni (ad esempio la distributività), valgono anche in F per le operazioni indotte.

Dimostrazione. La funzione e è ancora f stessa, perchè f è un idempotente e i è semplicemente l'inclusione di F in X. Dunque se vogliamo definire un'operazione $x \circ y$ su F in modo che e sia un omomorfismo, cioè $f(a*b) = f(a) \circ f(b)$, allora se $x = f(x)$ e $y = f(y)$ sono due punti fissi si deve avere

$$x \circ y = f(x) \circ f(y) = f(x * y)$$

e dunque l'operazione su F deve essere definita così. Viceversa, definendo in tal modo l'operazione su F, la proprietà di chiusura assicura che la proiezione $e\colon X \longrightarrow F$ è un omomorfismo. Infatti, per ogni $a, b \in X$ si ha: $f(a * b) = f[f(a) * f(b)] = f(a) \circ f(b)$. Inoltre si ha subito che $u' = f(u)$ è l'elemento neutro per l'operazione $\circ$ su F: se $x = f(x)$ è un punto fisso, allora $u' \circ x = f(u) \circ f(x) = f[f(u) * f(x)] = f(u * x) = f(x) = x$; similmente si ha $x \circ u' = x$. Non rimane che provare l'associatività; se $z = f(z)$ è un altro punto fisso, allora
$(x \circ y) \circ z = f[(x \circ y) * z] = f[f(x * y) * z] = f[f(x * y) * f(z)] = f[(x * y) * z] = f[x * (y * z)] = f[f(x) * f(y * z)] = x \circ (y \circ z)$.
Nello stesso modo si mostra che le ulteriori identità che possono essere soddisfatte dal monoide $(X, *, u)$ (ad esempio la commutatività), sono soddisfatte anche dal monoide $(F, \circ, u')$.

Viceversa, se su F esiste una operazione "$\circ$" per cui e è un omomorfismo, cioè $f(a * b) = f(a) \circ f(b)$, allora f *deve* soddisfare la proprietà di chiusura:
$f[f(a) * f(b)] = f(f(a)) \circ f(f(b)) = f(a) \circ f(b) = f(a * b)$.

Infine, supponiamo che su X ci sia un'altra operazione di monoide "$\Box$" per cui f soddisfa la proprietà di chiusura e chiamiamo "$\cdot$" l'operazione indotta su F (cioè $x \cdot y = f(x \Box y)$). Supponiamo che in X valga la distributività $a \Box (b * c) = (a \Box b) * (a \Box c)$; allora la stessa distributività vale in F per le operazioni indotte:
$x \cdot (y \circ z) = f(x \Box (y \circ z)) = f[f(x) \Box f(y * z)] = f[x \Box (y * z)] = f[(x \Box y) * (x \Box z)] = f[f(x \Box y) * f(x \Box z)] = (x \cdot y) \circ (x \cdot z)$. ■

Applichiamo questo teorema all'idempotente f su $\mathbf{N} \times \mathbf{N}$ descritto precedentemente. È facile verificare che f soddisfa la proprietà di chiusura rispetto alla somma 'puntuale' su $\mathbf{N} \times \mathbf{N}$ (esercizio che si svolge distinguendo i vari casi). Dunque, l'insieme dei punti fissi $\mathbf{Z}$ ha un'unica struttura di monoide commutativo per cui e è un omomorfismo, che tradizionalmente denotiamo ancora con il simbolo "$+$", perchè l'inclusione $i\colon \mathbf{N} \hookrightarrow \mathbf{Z}$ definita da $i(n) = \langle 0, n\rangle$ è un omomorfismo (esercizio). Il fatto che la costruzione di $\mathbf{Z}$ risolve il problema di aggiungere i *negativi* al monoide $(\mathbf{N}, +)$ è di immediata verifica: ogni elemento $x \in \mathbf{Z}$ ha un opposto $-x$, cioè un elemento $-x$ tale che $x + (-x) = 0$. In altre parole, $\mathbf{Z}$ è un *gruppo* rispetto alla somma, e sappiamo che dunque l'opposto è unico: l'opposto di $\langle 0, d\rangle$ è $\langle d, 0\rangle$ e viceversa.

La difficoltà è che l'idempotente f *non* è un operatore di chiusura rispetto al prodotto puntuale (si trovi un controesempio). È possibile però trovare un'altro prodotto su $\mathbf{N}\times\mathbf{N}$ per cui sia un monoide commutativo e che sia distributivo rispetto alla somma, rispetto al quale f è un operatore di chiusura (si veda l'esercizio 2.13.1). Tuttavia ora preferiamo mostrare che il prodotto in $\mathbf{Z}$ è unicamente determinato dalle condizioni di essere conservato dall'inclusione $i\colon \mathbf{N} \hookrightarrow \mathbf{Z}$ e dalla richiesta di essere associativo e *distributivo* rispetto alla somma. Infatti, la richiesta che i conservi il prodotto equivale a richiedere che il prodotto di 'positivi', cioè il prodotto di coppie $\langle 0, d\rangle$, $\langle 0, e\rangle$ sia definito come la coppia $\langle 0, de\rangle$. Poichè in un semi-anello in cui la somma è un gruppo, cioè in un *anello*, la distributività implica la *regola dei segni* (lemma 1.16.1), si vede allora che gli altri possibili prodotti sono determinati come
$\langle d, 0\rangle\langle 0, e\rangle = \langle de, 0\rangle$, $\langle 0, d\rangle\langle e, 0\rangle = \langle de, 0\rangle$,
$\langle d, 0\rangle\langle e, 0\rangle = \langle 0, de\rangle$.
Non resta che verificare che il prodotto così definito è in effetti associativo e distributivo rispetto alla somma. La commutatività di tale prodotto segue dalla commutatività del prodotto in $\mathbf{N}$.

2.8 Divisione di interi

Poichè nell'anello degli interi $\mathbf{Z}$ esiste la divisione, si ottengono altri esempi di anelli commutativi nel modo seguente. Ricordiamo dapprima che l'esistenza della divisione in $\mathbf{Z}$ significa che per ogni coppia di interi x e n, con $n > 0$, esiste un'*unica* coppia di interi q e r, detti rispettivamente 'quoziente' e 'resto', tali che

$$x = qn + r \quad \text{e} \quad 0 \leq r < n\ .$$

Infatti, se $x \geq 0$, allora siamo nel caso discusso per i naturali; se $x < 0$, allora $0 \leq (-x - 1)$ e dunque esistono q e r tali che $-x - 1 = nq + r$ e $0 \leq r < n$; dunque $x = n(-q - 1) + (n - r - 1)$ e $0 \leq (n - r - 1) < n$.

L'unicità del resto permette di definire, per ogni $n > 0$ fissato, una funzione

$$\mathbf{Z} \xrightarrow{\ r_n\ } \mathbf{Z}$$

come:

$$r_n(x) = \text{l'unico resto della divisione di } x \text{ per } n \ .$$

È immediato constatare che la funzione r_n è *idempotente*; denotiamo con $\mathbf{Z}_n$ l'insieme dei suoi punti fissi. $\mathbf{Z}_n$ risulta essere l'insieme *dei possibili resti della divisione per* n, cioè

$$\mathbf{Z}_n = \{0, 1, 2, \ldots, n-1\} \ .$$

È naturale a questo punto chiedersi se r_n soddisfa le proprietà di chiusura rispetto alla somma ed al prodotto in $\mathbf{Z}$, cioè se

$$r_n(r_n(x) + r_n(y)) = r_n(x+y) \quad , \quad r_n(r_n(x)r_n(y)) = r_n(xy) \ .$$

La risposta è affermativa per entrambe. Dimostriamo la prima, lasciando al lettore la dimostrazione della seconda. Dividiamo x e y per n:
$x = q_1 n + r_n(x)$, $y = q_1 n + r_n(y)$;
eseguiamo anche la divisione
$r_n(x) + r_n(y) = q_3 n + r_n(r_n(x) + r_n(y))$.
Sommando e sostituendo si ha:
$x + y = (q_1 + q_2)n + r_n(x) + r_n(y) = (q_1 + q_2 + q_3)n + r_n(r_n(x) + r_n(y))$.
Poichè $r_n(r_n(x) + r_n(y)) < n$, l'unicità del resto garantisce che tale quantità è il resto della divisione di $x + y$ per n. Per il teorema sugli operatori di chiusura si ha che $\mathbf{Z}_n$ è un *anello commutativo* e che la proiezione $\mathbf{Z} \longrightarrow \mathbf{Z}_n$ data dal prendere il resto della divisione per n è un *omomorfismo di anelli*, per le operazioni su $\mathbf{Z}_n$ che consistono nel prendere il resto della divisione per n delle operazioni omonime su $\mathbf{Z}$.

Per ogni n fissato, la relazione di equivalenza su $\mathbf{Z}$ data dal nucleo di equivalenza di r_n (per una definizione precisa di tali concetti si veda il paragrafo 6 del capitolo 3) si chiama "*congruenza modulo* n" e la si indica con

$$x \equiv y \pmod{n} \, .$$

Dunque $x \equiv y \pmod{n}$ se e solo se $x - y$ è un multiplo di n (esercizio).

2.9 Massimo Comune Divisore

Un'altra importante conseguenza della divisione tra interi è l'esistenza del *massimo comune divisore* di due interi. Diciamo che un intero x

divide un intero y (e scriviamo $x|y$) se esiste un intero k tale che $y = kx$. Il massimo comune divisore (M.C.D.) di due interi x e y è un intero $d = \text{M.C.D.}(x, y)$ tale che $d|x$ e $d|y$ e tale che, per ogni altro intero k, se $k|x$ e $k|y$, allora $k|d$. È chiaro che il M.C.D. è unico a meno del segno. Il seguente teorema dimostra l'esistenza del M.C.D. mediante un algoritmo per determinarlo.

Teorema 2.9.1 i) *Se x e y sono due interi non nulli allora esiste il massimo comune divisore $d = M.C.D.(x, y)$.*

ii) *Esistono due interi r e s tali che $d = rx + sy$.*

DIMOSTRAZIONE. Dimostreremo l'esistenza del massimo comune divisore fornendo un procedimento effettivo per calcolarlo, noto come "*algoritmo euclideo delle divisioni successive*". Possiamo supporre, senza alterare la generalità della dimostrazione, che $0 \leq y$ (perchè?). Si consideri la catena delle divisioni:

$x = qy + r, \qquad 0 \leq r < y;$
$y = q_1 r + r_1, \qquad 0 \leq r_1 < r;$
$r = q_2 r_1 + r_2, \qquad 0 \leq r_2 < r_1;$
. . . . . .

Poichè ogni resto è maggiore o uguale a zero e minore del precedente, *tale catena deve terminare*: per un certo n si deve avere $r_{n+1} = 0$, dunque:

. . . . . .
$r_{n-2} = q_n r_{n-1} + r_n, \qquad 0 \leq r_{n-1} < r_n;$
$r_{n-1} = q_{n+1} r_n.$

Questo fatto prova che il M.C.D. di x e y è proprio *l'ultimo resto non nullo* r_n. Infatti, per l'ultima uguaglianza, si ha che r_n divide r_{n-1}, quindi per la penultima uguaglianza divide anche r_{n-2} e, risalendo così la catena delle uguaglianze, si vede che divide anche y e x. Inoltre, se k divide x e y, allora divide anche $r = x - qy$, dunque divide anche $r_1 = y - q_1 r$ e, discendendo così la catena delle uguaglianze, si vede che divide anche r_{n-1}. Dunque r_n è il M.C.D. di x e y.

ii) Per dimostrare che esistono due numeri r e s tali che $d = \text{M.C.D.}(x, y) = rx + sy$, esprimiamo il M.C.D. di x e y come ultimo resto non nullo nell'algoritmo euclideo delle divisioni successive. La penultima di tali uguaglianze fornisce:

$r_n = r_{n-2} - q_n r_{n-1}$;
la precedente uguaglianza ($r_{n-3} = q_{n-1} r_{n-2} + r_{n-1}$) fornisce r_{n-1} come somma di prodotti (= combinazione lineare) di r_{n-3} e r_{n-2} che, sostituito nella precedente espressione per r_n, fornisce, dopo aver sommato i termini simili, una espressione di r_n come combinazione lineare di r_{n-3} e r_{n-2}; così continuando a risalire la catena delle uguaglianze, si ottiene alla fine una espressione di r_n come combinazione lineare di x e y. ■

Vediamo alcune applicazioni del precedente teorema sul M.C.D ai gruppi $\mathbf{Z}_n$. Ricordiamo che un numero intero p è *primo* se gli unici suoi divisori sono, a meno del segno, 1 e p; ricordiamo anche che due numeri interi x e y si dicono *primi relativi* se M.C.D.$(x, y) = 1$.

Teorema 2.9.2 . *Un elemento non nullo di* $\mathbf{Z}_n$ *è invertibile se e solo se* M.C.D.$(x, n) = 1$.

DIMOSTRAZIONE. Se M.C.D.$(x, n) = 1$, allora per il teorema sul M.C.D., esistono due interi r e s tali che $1 = ax + bn$. Dunque, prendendo i resti modulo n di entrambi i membri di tale uguaglianza e tenendo conto che la funzione r_n "prendere in resto modulo n" è un operatore di chiusura rispetto alla somma e al prodotto, si ha:
$1 = r_n(1) = r_n(ax+bn) = r_n[r_n(ax)+r_n(bn)] = r_n(ax) = r_n[r_n(a)r_n(x)] = r_n[r_n(a)x]$.
Dunque $r_n(a)$ è l'inverso di x in $\mathbf{Z}_n$.

Viceversa, supponiamo che $x \in \mathbf{Z}_n$ sia invertibile, cioè che esista $p \in \mathbf{Z}_n$ tale che $px = 1$ in $\mathbf{Z}_n$; ciò significa che il resto della divisione di px per n è 1, dunque che $px = hn + 1$; ma allora M.C.D.$(x, n) = 1$. ■

Come ulteriore applicazione del teorema sul M.C.D. diamo un metodo per trovare tutte le soluzioni intere di una equazione del tipo

$$ax + by = c\,,$$

dove a, b e c sono coefficienti interi ("*equazione diofantea lineare*"). Cominciamo con l'osservare che tale equazione ha una soluzione se e solo se $d =$ M.C.D.(a, b) divide il termine noto c. Infatti, poichè d divide a e d divide b, esistono h e k tali che $a = hd$ e $b = kd$; dunque se l'equazione ha una soluzione x, y, allora $c = ax + by = hdx + kdy = (hx + ky)d$; dunque d divide c. Viceversa, se d divide c, dunque $c = dt$, le soluzioni

dell'equazione sono le stesse di quella che si ottiene dividendo per d:

$$hx + ky = t\,;$$

poichè h e k sono primi relativi, il teorema sul M.C.D. assicura che esistono due interi r e s tali che $hr + ks = 1$; moltiplicando per t si ha: $h(rt)+k(st) = t$; dunque $x = rt$ e $y = st$ è una soluzione dell'equazione ridotta, quindi anche di quella di partenza. Tutte le altre soluzioni si ottengono aggiungendo ad una soluzione particolare le soluzioni della equazione omogenea associata:

$$hx + ky = 0\,.$$

Infatti, se x_0, y_0 è una soluzione dell'equazione omogenea, allora $x + x_0, y + y_0$ è ancora una soluzione dell'equazione (ridotta), poichè $h(x + x_0)+k(y+y_0) = (hx+ky)+(hx_0+ky_0) = t+0 = t$. Inoltre, se x', y' è una soluzione dell'equazione (ridotta), allora $x - x', y - y'$ è una soluzione dell'equazione omogenea. Poichè tutte le soluzioni dell'equazione omogenea sono evidentemente: $x_0 = mb, y_0 = -ma$, per $m \in \mathbf{Z}$, allora tutte le soluzioni della equazione data sono $x = rt + mb, y = st - ma$, per $m \in \mathbf{Z}$.

Una applicazione della precedente discussione è il "*teorema cinese del resto*": se p e q sono primi relativi, allora la coppia di congruenze:

$$x \equiv a \pmod{p}, \quad x \equiv b \pmod{q}\,,$$

ha una sola soluzione (mod pq). Infatti, i numeri x che soddisfano contemporaneamente le due congruenze date sono quelli per cui esistono h e k tali che $x = hp + a = kq + b$. Basta dunque risolvere l'equazione $hp - kq = b - a$ che, per quando discusso precedentemente ha sicuramente soluzioni, perchè M.C.D.$(p, q) = 1$. Se r e s sono due numeri per cui $1 = rp + sq$, le soluzioni sono:
$h = r(b - a) + mq, k = s(a - b) + mp$, per $m \in \mathbf{Z}$;
dunque le soluzioni delle congruenze date sono:
$x = r(b - a)p + mpq + a = s(a - b)q + mpq + b$, per $m \in \mathbf{Z}$
cioè definiscono un elemento di $\mathbf{Z}_{pq}$ tale che $r_p(x) = a$ e $r_q(x) = b$. Una conseguenza di tale fatto è la seguente. Sia

$$\alpha\colon \mathbf{Z}_{pq} \longrightarrow \mathbf{Z}_p \times \mathbf{Z}_q$$

la funzione che ad ogni intero $0 \leq x < pq$ associa la coppia $\alpha(x) = \langle r_p(x), r_q(x)\rangle$. Il teorema cinese del resto assicura che tale funzione è suriettiva e quindi una corrispondenza biunivoca, perchè entrambi gli insiemi hanno la stessa cardinalità. È anche facile vedere direttamente che α conserva la somma, il prodotto e 1 e dunque che è un *isomorfismo di anelli*, pur di definire la struttura di anello sul prodotto $\mathbf{Z}_p \times \mathbf{Z}_q$ componente per componente (prodotto di anelli). In particolare, questo fatto ci assicura che α definisce un isomorfismo tra gli anelli $\mathbf{Z}_{pq}^*$ e $\mathbf{Z}_p^* \times \mathbf{Z}_q^*$. È un fatto generale che se $f: A \longrightarrow B$ è un isomorfismo di anelli, esso induce un isomorfismo tra i gruppi A^* e B^* degli elementi invertibili di A e B (esercizio).

Per finire non possiamo a questo punto non menzionare una famosa funzione, la funzione $\phi: \mathbf{N} - \{0\} \longrightarrow \mathbf{N} - \{0\}$ che ad ogni naturale $n \neq 0$ associa il numero $\phi(n)$ dei naturali minori di n e primi con n. Alcuni suoi valori sono perciò:$\phi(1) = 1$, $\phi(2) = 1$, $\phi(3) = 2$, $\phi(4) = 2$, $\phi(5) = 4$ e, in generale, se p è un numero primo, $\phi(p) = p - 1$. La funzione ϕ è detta "*funzione di Eulero*" e per il teorema precedente, il numero $\phi(n)$ coincide con l'ordine del gruppo $\mathbf{Z}_n^*$ degli elementi invertibili di $\mathbf{Z}_n$. Il teorema cinese del resto garantisce che se p e q sono coprimi, allora $\phi(pq) = \phi(p)\phi(q)$. Quindi, il *teorema fondamentale dell'aritmetica*, che dice che ogni naturale $n > 1$ si fattorizza in modo unico (a meno dell'ordine) come prodotto

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k},$$

dove p_i sono primi e α_i sono naturali non nulli, ci permette di dare una formula esplicita per $\phi(n)$, pur di conoscere una formula esplicita per $\phi(p^\alpha)$, dove p è un primo e α è un naturale non nullo.

2.10 Esercizi

1. Si provi che se n e a sono coprimi e se $n|ab$, allora $n|b$.

2. Si provi che se p è primo e $p|ab$, allora $p|a$ o $p|b$.

3. Se p è un primo, si provi che il numero dei numeri compresi tra 1 e p^α che *non* sono coprimi con p^α è $p^{\alpha-1}$. Si deduca il valore di

$\phi(p^\alpha)$ e, usando il teorema cinese del resto insieme al teorema fondamentale dell'aritmetica, si dia una formula esplicita per $\phi(n)$.

2.11 Numeri Primi

Il teorema fondamentale dell'aritmetica, noto già ai Greci, ha numerose conseguenze oltre a quello del calcolo esplicito della funzione ϕ. Almeno due di queste sono argomenti noti fino dalla antichità, che per la loro semplicità e universalità dovrebbero appartenere al bagaglio di conoscenze di ogni persona acculturata, anche se non è un matematico professionista.

Teorema 2.11.1 (Infinità dei primi (Euclide).) *Esistono infiniti numeri primi.*

DIMOSTRAZIONE. Per assurdo: supponiamo che il loro numero sia finito e sia p "*l'ultimo numero primo*". Consideriamo il numero

$$p! + 1 .$$

Poichè $p! + 1$ ha una (unica) fattorizzazione in potenze di primi, sia q un suo fattore primo. Poichè p è l'ultimo primo, $p!$ dovrà contenere q come fattore, dunque il resto della divisione di $p! + 1$ per q è 1, dato che $1 < q$. D'altra parte il resto della divisione di $p! + 1$ per q è 0, perchè q è un suo fattore, ciò che contraddice l'unicità del resto della divisione. ■.

Un'altro semplicissimo argomento ancora basato sul teorema della fattorizzazione unica in primi è la "*non razionalità della radice quadrata di* 2" e lo richiameremo in seguito. Questi esempi dovrebbero essere sufficienti a convincerci dell'interesse di dare una dimostrazione di tale teorema. Il lettore interessato può trovare una dimostrazione elementare per induzione ad esempio nel libro di Herstein (Algebra, Editori Riuniti, 1982). Un'altra dimostrazione più generale, ma non costruttiva, verrà data in seguito nella teoria degli anelli ad ideali principali (si veda ad esempio il libro di Birkhoff - MacLane, "Algebra", Editori Riuniti, Cap. IV, Teorema 26).

2.12 I Numeri Razionali

L'ultimo esempio di costruzione della stessa natura delle precedenti che vogliamo esaminare è la costruzione dei *numeri razionali* $\mathbf{Q}$. La costruzione di $\mathbf{Q}$ risponde al desiderio di aggiungere ai numeri interi *non nulli* l'inverso rispetto al prodotto, cioè di aggiungere ad ogni intero non nullo x un (unico) numero x^{-1} tale che $x^{-1}x = 1$ (ricordiamo che un anello commutativo con tale proprietà è detto '*campo*'). L'idea è che un numero razionale è una espressione formale xy^{-1}, dove y è non nullo, che interpretiamo come soluzione dell'equazione $yz = x$ nell'incognita z, che tradizionalmente si denota con la scrittura come '*frazione*' $\frac{x}{y}$. Essendo tale scrittura una scrittura formale, possiamo considerarla semplicemente come una coppia di interi $\langle x, y\rangle$, con y non nullo. Dunque, possiamo partire dall'insieme

$$\mathbf{Z} \times (\mathbf{Z} - \{0\})$$

delle coppie ordinate di interi la cui seconda componente è non nulla. Tuttavia, tra tutte tali coppie hanno un ruolo essenziale quelle che *non hanno divisori comuni*. Questo ci permette di descrivere $\mathbf{Q}$ come l'insieme dei punti fissi di un idempotente su $\mathbf{Z} \times (\mathbf{Z} - \{0\})$ nel modo seguente.

Ricordiamo dapprima che la divisione in $\mathbf{Z}$ permette di dimostrare l'esistenza del *massimo comune divisore* (x, y) di ogni coppia di interi non entrambi nulli x e y (ad esempio mediante l'algoritmo euclideo delle divisioni successive), cioè di un intero d che divide sia x sia y e tale che sia diviso da ogni divisore comune di x e y. Poichè il massimo comune divisore è *unico a meno del segno*, conveniamo di scegliere quello positivo. Possiamo dunque definire una funzione

$$\mathbf{Z} \times (\mathbf{Z} - \{0\}) \xrightarrow{d} \mathbf{Z} \times (\mathbf{Z} - \{0\})$$

mediante

$$d(x, y) = \langle \frac{x}{(x, y)}, \frac{y}{(x, y)} \rangle$$

dove il simbolo di frazione indica il quoziente in $\mathbf{Z}$ della divisione il cui resto, nelle nostre ipotesi, è nullo. È immediato vedere che d è un idempotente; definiamo $\mathbf{Q}$ come l'insieme dei punti fissi di d. $\mathbf{Q}$ così

definito coincide con i numeri razionali come li conosciamo dalle scuole medie, poichè è isomorfo alle frazioni $\frac{x}{y}$ 'ridotte ai minimi termini', cioè alle coppie di interi $\langle x, y\rangle$ con y non nullo e con $(x, y) = 1$ (si dice che x e y sono 'coprimi'; si osservi che l'unico numero coprimo con 0 è 1, perchè se $y \neq 0$, allora $(0, y) = y$).

Rimane il problema di spiegare le usuali operazioni su $\mathbf{Q}$. Ora, $\mathbf{Z} \times (\mathbf{Z} - \{0\})$ è un monoide commutativo rispetto al prodotto puntuale, e non è difficile vedere che d soddisfa la proprietà di chiusura rispetto al prodotto puntuale (si veda l'esercizio 3). Dunque, per il teorema sugli operatori di chiusura, si ha che $\mathbf{Q}$ è un monoide commutativo rispetto al prodotto

$$\langle x, y\rangle * \langle r, s\rangle = d(\langle x, y\rangle\langle r, s\rangle) = d(\langle xr, ys\rangle) \ ,$$

che è nient'altro che la definizione che si usa in pratica: si fa il prodotto dei numeratori e dei denominatori e poi si riduce ai minimi termini. Si osservi che la funzione

$$i\colon \mathbf{Z} \hookrightarrow \mathbf{Q}$$

definita da $i(x) = \langle x, 1\rangle$ è un omomorfismo iniettivo di monoidi, cioè i è iniettiva e $\langle x, 1\rangle * \langle y, 1\rangle = \langle xy, 1\rangle$. Perciò il prodotto in $\mathbf{Q}$ 'estende' in prodotto in $\mathbf{N}$ e per questa ragione ometteremo la notazione '$*$' per il prodotto in $\mathbf{Q}$.

La ragione per cui è legittima la convenzione di scrivere una coppia (x, y) di interi coprimi con y non nullo nella forma di frazione $\frac{x}{y}$, è la seguente: *ogni razionale* $(x, y) \neq \langle 0, 1\rangle$ *è invertibile*: $\langle x, y\rangle\langle y, x\rangle = d(xy, xy) = \langle 1, 1\rangle$, che è l'elemento neutro del prodotto in $\mathbf{Q}$; dunque, $\langle x, y\rangle^{-1} = \langle y, x\rangle$. In particolare si ha

$$\langle x, y\rangle = \langle x, 1\rangle\langle 1, y\rangle = \langle x, 1\rangle\langle y, 1\rangle^{-1} \ ,$$

ciò che giustifica la notazione

$$\langle x, y\rangle = \frac{x}{y} \ .$$

Infine osserviamo che l'unico elemento non invertibile $\langle 0, 1\rangle$ di $\mathbf{Q}$, che denoteremo ancora con 0 perchè è $i(0)$, ha la proprietà $0\langle x, y\rangle = 0$, per ogni $\langle x, y\rangle \in \mathbf{Q}$.

Rimane da spiegare la somma di razionali. Osserviamo che la somma puntuale su $\mathbf{Z}\times(\mathbf{Z}-\{0\})$ è una operazione rispetto alla quale il prodotto puntuale è distributivo, ma che non ha elemento neutro, perchè dovrebbe essere $\langle 0,0\rangle$. Inoltre, l'idempotente d *non* soddisfa la proprietà di chiusura rispetto alla somma puntuale (si trovi un controesempio). Siamo in una situazione analoga ma opposta a quella descritta per la costruzione degli interi $\mathbf{Z}$. Tuttavia vediamo che anche in tal caso la somma è determinata dalla richiesta che l'inclusione $i\colon \mathbf{Z} \hookrightarrow \mathbf{Q}$ sia un omomorfismo di anelli e dalla richiesta che il prodotto sia distributivo rispetto a tale somma. Infatti, supponiamo che una tale somma esista e denotiamola semplicemente con '+'. La prima condizione equivale alla condizione $\langle x,1\rangle + \langle y,1\rangle = \langle x+y,1\rangle$. Usando la seconda condizione, calcoliamo il prodotto:
$(\langle x,y\rangle + \langle r,s\rangle)\langle ys,1\rangle = \langle x,y\rangle\langle ys,1\rangle + \langle r,s\rangle\langle ys,1\rangle =$
$= d(xys,y) + d(rys,s) = \langle xs,1\rangle + \langle ry,1\rangle = \langle xs+ry,1\rangle.$
Dunque $(\langle x,y\rangle + \langle r,s\rangle)\langle ys,1\rangle = \langle xs+ry,1\rangle$. Moltiplicando entrambi i membri per l'inverso di $\langle ys,1\rangle$ che è $\langle 1,ys\rangle$, si ottiene
$\langle x,y\rangle + \langle r,s\rangle = \langle xs+ry,1\rangle\langle 1,ys\rangle = d(xs+ry,ys),$
che è la formula usuale per la somma di frazioni ridotte ai minimi termini. Dunque, se la somma con le proprietà prescritte esiste, deve essere data dalla formula precedente. Si tratta di dimostrare che, definendo la somma come sopra, essa ha le proprietà prescritte (si veda anche il seguente esercizio 4).

2.13 Esercizi

1. Dimostrare che l'operazione '$\Box$' su $\mathbf{N}\times\mathbf{N}$ definita da

$$\langle n,m\rangle \Box \langle r,s\rangle = \langle ns+mr, nr+ms\rangle$$

è una operazione associativa, commutativa, per cui $\langle 0,1\rangle$ è l'elemento neutro e che è distributiva rispetto alla somma puntuale; dunque essa induce una struttura di semi-anello su $\mathbf{N}\times\mathbf{N}$ diversa da quella in cui il prodotto di coppie è definito puntualmente. Dimostrare anche che l'idempotente f definito all'inizio del paragrafo è un omomorfismo rispetto a questa operazione e

che quindi in particolare soddisfa la proprietà di chiusura rispetto a '$\square$'. Mostrare infine che l'operazione indotta sull'insieme $\mathbf{Z}$ dei punti fissi coincide con l'usuale prodotto di interi.

2. Dimostrare che la funzione resto r_n ha la proprietà di chiusura rispetto al prodotto.

3. Dimostrare che l'idempotente d su $\mathbf{Z} \times (\mathbf{Z} - \{0\})$ definito precedentemente soddisfa la proprietà di chiusura rispetto al prodotto puntuale. (Suggerimento: se $h = (x, y)$ ed $e = (r, s)$, allora $x = x_1 h$, $y = y_1 h$, $r = r_1 e$, $s = s_1 e$; dunque

   $(\langle x, y\rangle\langle r, s\rangle) = d(xr, ys) = \langle \frac{xr}{(xr,ys)}, \frac{ys}{(xr,ys)}\rangle =$

   $= \langle \frac{x_1 r_1}{(x_1 r_1, y_1 s_1)}, \frac{y_1 s_1}{(x_1 r_1, y_1 s_1)}\rangle$

   perchè

   $(xr, ys) = (x_1 r_1 he, y_1 s_1 he) = he(x_1 r_1, y_1 s_1)$;

   eseguendo calcoli analoghi sull'altro membro dell'equazione che esprime la proprietà di chiusura di d si arriva allo stesso risultato.)

4. Dimostrare che l'operazione $\langle x, y\rangle + \langle r, s\rangle = \langle xs + yr, ys\rangle$ sull'insieme $\mathbf{Z} \times (\mathbf{Z} - \{0\})$ è una operazione associativa, commutativa, con elemento neutro dato da $\langle 0, 1\rangle$ e per cui il prodotto puntuale è distributivo. Dimostrare che d soddisfa la proprietà di chiusura rispetto a tale somma e che l'operazione perciò indotta sui razionali coincide con la somma di razionali.

5. Dimostrare che le due costruzioni di $\mathbf{Z}$ e $\mathbf{Q}$ si possono eseguire nell'ordine opposto a quello presentato in questo paragrafo: partendo da $\mathbf{N}$ si può prima considerare l'idempotente d ma definito solo su $\mathbf{N} \times (\mathbf{N} - \{0\})$, i cui punti fissi sono i razionali non negativi $\mathbf{Q}^{\geq 0}$. Poi, osservando che l'ordine in $\mathbf{Q}^{\geq 0}$ è definibile come in $\mathbf{N}$ e che per esso vale la tricotomia, si può definire su $\mathbf{Q}^{\geq 0} \times \mathbf{Q}^{\geq 0}$ un idempotente f i cui punti fissi sono tutti i razionali $\mathbf{Q}$ in maniera analoga a quanto fatto per $\mathbf{N} \times \mathbf{N}$.

6. Si costruisca una corrispondenza biunivoca

$$\gamma\colon \mathbf{Q}^{\geq 1} \longrightarrow W(\mathbf{N}^{>0})$$

dall'insieme dei razionali maggiori o uguali a 1 al semigruppo delle parole sui naturali non nulli. (Suggerimento: se $x = \langle m, n\rangle$ è un razionale maggiore o uguale a 1, si consideri la sua "parte intera" $[x]$, che altro non è che il quoziente della divisione $m = n[x] + r_1$ $(0 \leq r_1 < n)$; allora $x_1 = \langle n, r_1\rangle$ è un razionale maggiore o uguale a 1 per cui

$$x = [x] + \frac{1}{x_1}\,.$$

È chiaro che questo processo può essere iterato un numero finito di volte, poichè ad un certo punto si deve ottenere un resto nullo. Questa procedura definisce in modo univoco una parola di naturali non nulli

$$w(x) = [x][x_1]\ldots[x_k]\,,$$

che si deve mostrare essere una corrispondenza biunivoca. La funzione inversa è detta "frazione continua finita").

7. Costruzione dei reali non negativi estesi come spezzamento di un idempotente. Un '*filtro*' F su $\mathbf{Q}^{\geq 0}$ è un sottoinsieme $F \subseteq \mathbf{Q}^{\geq 0}$ di $\mathbf{Q}^{\geq 0}$ con la proprietà

   $x \in F$ e $x \leq y \quad \Rightarrow \quad y \in F$.

   Sia $\mathcal{F}(\mathbf{Q}^{\geq 0})$ l'insieme dei filtri su $\mathbf{Q}^{\geq 0}$. Consideriamo la funzione

   $$\mathcal{F}(\mathbf{Q}^{\geq 0}) \xrightarrow{D} \mathcal{F}(\mathbf{Q}^{\geq 0})$$

   definita da:

   $$D(F) = \{q \in \mathbf{Q}^{\geq 0} \mid x > q \;\Rightarrow\; x \in F\}\,.$$

   Si provi che D è un idempotente su $\mathbf{Q}^{\geq 0}$ e che l'insieme dei suoi punti fissi è l'insieme $\overline{\mathbf{R}^{\geq 0}}$ dei numeri reali non negativi con l'aggiunta del punto all'infinito. Quali operazioni su $\mathbf{Q}^{\geq 0}$ si trasportano su $\mathcal{F}(\mathbf{Q}^{\geq 0})$ e per quali di esse l'idempotente D soddisfa la proprietà di chiusura?

8. Si provi che la stessa funzione w definita nell'esercizio 6 fornisce una corrispondenza biunivoca

   $$\gamma\colon \mathbf{I}^{>1} \longrightarrow \mathbf{N}^{>0^{\mathbf{N}}}$$

dall'insieme $\mathbf{I}^{>1} = \mathbf{R}^{>1} - \mathbf{Q}^{>1}$ degli irrazionali maggiori di 1 all'insieme delle successioni di naturali maggiori di 0. La funzione inversa è detta "frazione continua".

Capitolo 3

Omomorfismi

3.1 Immagini dirette ed inverse

Sia $f: X \longrightarrow Y$ una funzione e sia $V \subseteq Y$ un sottoinsieme di Y. L'insieme:

$$f^*(V) = \{x \in X | f(x) \in V\}$$

è un sottoinsieme di X detto *immagine inversa di V nella funzione f*. Se V è costituto da un solo elemento $y \in Y$, allora l'immagine inversa $f^*(\{y\})$ di V viene denotata semplicemente con $f^*(y)$ e viene detta *controimmagine di y*; dunque:

$$f^*(y) = \{x \in X | f(x) = y\}.$$

Se $U \subseteq X$ è un sottoinsieme di X, l'insieme

$$f_*(U) = \{y \in Y | \text{esiste } x \in U: f(x) = y\}$$

è un sottoinsieme di Y detto *immagine* (*diretta*) di U nella funzione f. In particolare il sottoinsieme di Y dato da

$$f_*(X) = \{y \in Y | \text{esiste } x \in X: f(x) = y\}$$

è detto *immagine della funzione f* e denotato con $\text{Im}(f)$. Ricordando che cosa si intende per funzione suriettiva, si vede che f è suriettiva se e solo se $\text{Im}(f) = Y$. In particolare, ogni funzione $f: X \longrightarrow Y$

diventa una funzione suriettiva quando la si considera come funzione $f\colon X \longrightarrow \mathrm{Im}(f)$.

Le precedenti definizioni di immagine inversa e di immagine diretta forniscono due funzioni

$$f^*\colon \mathbf{P}Y \longrightarrow \mathbf{P}X \quad , \quad f_*\colon \mathbf{P}X \longrightarrow \mathbf{P}Y$$

che soddisfano le seguenti identità:

1. $f^*(V_1 \cup V_2) = f^*(V_1) \cup f^*(V_2)$
2. $f^*(V_1 \cap V_2) = f^*(V_1) \cap f^*(V_2)$
3. $f_*(U_1 \cup U_2) = f_*(U_1) \cup f_*(U_2)$
4. $U \subseteq f^*(f_*(U))$ e $f_*(f^*(V)) \subseteq V$,

per ogni coppia di sottoinsiemi U_1, U_2 di X e V_1, V_2 di Y. A titolo di esempio dimostriamo la seconda uguaglianza. Poichè dobbiamo dimostrare che due sottoinsiemi di X sono uguali, dobbiamo usare il principio di estensionalità, cioè dobbiamo dimostrare che hanno gli stessi elementi: $x \in f^*(V_1 \cap V_2)$ se e solo se $f(x) \in V_1 \cap V_2$, cioè se e solo se $f(x) \in V_1$ e $f(x) \in V_2$; d'altra parte, $x \in f^*(V_1) \cap f^*(V_2)$ se e solo se $x \in f^*(V_1)$ e $x \in f^*(V_2)$, cioè se e solo se $f(x) \in V_1$ e $f(x) \in V_2$. Dunque i sottoinsiemi $f^*(V_1 \cap V_2)$ e $f^*(V_1) \cap f^*(V_2)$ hanno gli stessi elementi, perciò sono uguali.

Quando A e B sono dotati di una stessa struttura algebrica e f è un *omomorfismo*, è importante osservare che gli "*operatori*" f_* e f^* si restringono alle sottostrutture, cioè se $U \subseteq X$ e $V \subseteq Y$ sono *sottostrutture* di quelle su X e Y, allora l'immagine diretta $f_*(V)$ e l'immagine inversa $f^*(U)$ sono ancora sottostrutture di quelle di Y e di X. Un caso particolare è in realtà già stato dimostrato alla fine del capitolo 1, quando si è parlato di immagine e di nucleo (quest'ultimo nel caso dei monoidi). La dimostrazione generale non è molto diversa da quella data per l'immagine ed il nucleo e la lasciamo per esercizio.

3.2 Esercizi

1. Si provi che se $f: X \longrightarrow Y$ è una funzione e se U_1 e U_2 sono due sottoinsiemi di X, allora

$$f_*(U_1 \cap U_2) \subseteq f_*(U_1) \cap f_*(U_2)$$

   e si mostri con un controesempio che in generale *non* si ha l'uguaglianza.

2. Si provi che una funzione $f: X \longrightarrow Y$ è iniettiva se e solo se per ogni sottoinsieme $U \subseteq X$ si ha $f^*(f_*(U)) = U$ e che è suriettiva se e solo se per ogni sottoinsieme $V \subseteq Y$ si ha $f_*(f^*(V)) = V$.

3. Si provi che gli operatori f_* e f^* soddisfano la seguente proprietà: se $U \subseteq X$ e $V \subseteq Y$ sono due sottoinsiemi di X e Y, allora $f_*(U) \subseteq V$ se e solo se $U \subseteq f^*(V)$.

4. Si provino le identità 1, 2, 3 e 4 solo a partire dalla precedente proprietà degli operatori f_* e f^* e dal fatto che essi conservino l'ordine.

5. Se $(X, \circ, \exp)$ e $(Y, \cdot, u)$ sono monoidi e se $fX: \longrightarrow Y$ è un omomorfismo di monoidi, si provi che per ogni sottomonoide $U \subseteq X$ e $V \subseteq Y$ le immagini dirette e inverse $f_*(U)$ e $f^*(V)$ sono sottomonoidi rispettivamente del codominio e del dominio di f. Si enunci e si dimostri l'analogo risultato per i gruppi e per gli anelli.

3.3 Funzioni suriettive

Mostriamo ora come si possa determinare il *numero* $E(n, m)$ *delle funzioni suriettive* $f: [n] \longrightarrow [m]$ semplicemente usando il fatto che ogni funzione $f: X \longrightarrow Y$ si può analizzare come composizione di una funzione suriettiva e di una iniettiva prima considerando f come funzione da X alla propria immagine $\mathrm{Im}(X)$ e poi considerando l'inclusione di $\mathrm{Im}(X)$ in Y. Osserviamo dapprima che

1. $E(n, m) = 0$, se $n < m$;

2. $E(n,0) = 0$ e $E(n,1) = 1$, se $n > 0$;

3. $E(n,n) = n!$.

Dato che una funzione $f\colon [n] \longrightarrow [m]$ può essere considerata come una funzione suriettiva sulla propria immagine, si ha che il numero delle funzioni $f\colon [n] \longrightarrow [m]$ che hanno per immagine un sottoinsieme di $[m]$ con k elementi è dato da

$$\binom{m}{k} E(n,k),$$

cioè dal numero $E(n,k)$ delle funzioni suriettive $f\colon [n] \longrightarrow [k]$ per il numero $\binom{m}{k}$ dei sottoinsiemi di $[m]$ con k elementi. Infine, l'insieme delle funzioni $[n] \longrightarrow [m]$ può essere descritto come la somma dell'insieme delle funzioni $[n] \longrightarrow [m]$ che hanno per immagine un insieme con un elemento, più quello delle funzioni $[n] \longrightarrow [m]$ che hanno per immagine un insieme con due elementi ecc., si ha che il numero m^n di tutte le funzioni $f\colon [n] \longrightarrow [m]$ può essere espresso come:

$$\boxed{m^n = \textstyle\sum_{k=1}^{m} \binom{m}{k} E(n,k)}.$$

Fissato n, facendo variare m tra 1 e n si ottengono n equazioni lineari nelle n incognite $x_i = E(n,i)$; si ha cioè un sistema di n equazioni in n incognite:

$$\begin{array}{rcl}
1^n & = & \binom{1}{1} x_1 \\
2^n & = & \binom{2}{1} x_1 + \binom{2}{2} x_2 \\
3^n & = & \binom{3}{1} x_1 + \binom{3}{2} x_2 + \binom{3}{3} x_3 \\
\ldots & \ldots & \ldots \\
n^n & = & \binom{n}{1} x_1 + \binom{n}{2} x_2 + \binom{n}{3} x_3 + \ldots + \binom{n}{n} x_n
\end{array}$$

Si osservi che la matrice $A = \left[\binom{i}{k}\right]$ di questo sistema è una matrice triangolare bassa i cui elementi sulla diagonale sono tutti 1; dunque A ha determinante 1. Perciò A è invertibile e non è difficile mostrare che la matrice inversa A^{-1} di A è:

$$A^{-1} = \left[(-1)^{i+k}\binom{i}{k}\right].$$

Infatti, detto c_{rs} il generico elemento della matrice prodotto $A \cdot A^{-1}$, si ha:

$c_{rs} = \sum_{i=1}^{n}(-1)^{s+i}\binom{r}{i}\binom{i}{s} = \sum_{i=1}^{n}(-1)^{s+i}\binom{r}{s}\binom{r-s}{i-s} =$

$= \binom{r}{s}\sum_{i=s}^{r}(-1)^{s+i}\binom{r-s}{i-s} = (\text{posto } k = r-s)$

$= \binom{r}{s}\sum_{i=0}^{k}(-1)^{i}\binom{k}{i};$

ora usando la formula del binomio di Newton si ha:

$0 = (1-1)^k = \sum_{i=0}^{k}(-1)^{i}\binom{k}{i}$, se $k \neq 0$, cioè se $r \neq s$;

dunque $c_{rs} = 0$ se $r \neq s$, mentre $c_{rr} = 1$; perciò $A \cdot A^{-1}$ è la matrice identica. Moltiplicando A^{-1} per il vettore colonna dei termini noti si ha che le soluzioni del sistema sono date dalla formula:

$$\boxed{x_i = E(n,i) = \sum_{k=1}^{n}(-1)^{i+k}\binom{i}{k}k^n}.$$

3.4 Partizioni

Ricordiamo che una *partizione* $\mathcal{U} = \{\mathcal{U}_\rangle\}_{\rangle\in\mathcal{I}}$ di un insieme X è una famiglia di sottoinsiemi non vuoti U_i di X tale che ogni elemento di X sta in uno ed un solo sottoinsieme della famiglia. Gli insiemi U_i sono detti "*classi*" o "*regioni*" della partizione.

Poniamoci ora il problema di contare il numero delle partizioni di $[n]$ in m classi e ragioniamo come segue. Osserviamo dapprima che se $f\colon X \longrightarrow I$ è una funzione suriettiva, allora la famiglia $\mathcal{U}_\{ = \{\mathcal{U}_\rangle\}_{\rangle\in\mathcal{I}}$ definita da

$$U_i = f^*(i) = \{x \in [n] | f(x) = i\}$$

è una partizione di X e, se I è finito di cardinalità m, è costituita da m classi *poichè f è suriettiva* e dunque ciascun insieme U_i non è vuoto. Viceversa, se $\mathcal{U} = \{\mathcal{U}_\rangle\}_{\rangle \in \mathcal{I}}$ è una partizione di X, si può costruire una funzione *suriettiva* $f_{\mathcal{U}}: X \longrightarrow I$ definendo $f_{\mathcal{U}}(x) = i$ se e solo se $x \in U_i$. Il fatto che $\mathcal{U}$ sia una partizione equivale al fatto che tale definizione dà proprio una funzione suriettiva. Inoltre, è facile mostrare che data un'altra funzione suriettiva $g: X \longrightarrow J$, si ha che $\mathcal{U}_\{ = \mathcal{U}_\}$ se e solo se esiste un (unico) isomorfismo $\sigma: I \longrightarrow J$ tale che $\sigma f f = g$.

Nel caso di $X = [n]$ e $I = [m]$, poichè partendo dalla stessa partizione di $[n]$ si possono costruire in questo modo $m!$ funzioni suriettive $f: [n] \longrightarrow [m]$ semplicemente permutando l'ordine degli indici $i \in [m]$, si ha che il numero $E(n, m)$ delle funzioni suriettive $f: [n] \longrightarrow [m]$ ed il numero $S(n, m)$ delle partizioni di $[n]$ in m classi sono legati dalla relazione:

$$\boxed{E(n, m) = m!S(n, m)}.$$

La formula per $E(n, m)$ fornisce dunque:

$$S(n, m) = \sum_{k=1}^{m} (-1)^{m+k} \binom{m}{k} \frac{k^n}{m!}.$$

I numeri $S(n, m)$ sono chiamati *numeri di Stirling di seconda specie.*

Diamo ora un metodo *induttivo* per generare i numeri $S(n, m)$.

Lemma 3.4.1 *I numeri $S(n, m)$ soddisfano le seguenti relazioni:*

1. $S(n, n) = 1$, *se* $0 \leq n$;
2. $S(n, 0) = 0$, *se* $0 < n$;
3. $S(n, m) = 0$, *se* $n < m$;
4. $S(n + 1, m) = S(n, m - 1) + mS(n, m)$.

DIMOSTRAZIONE. Solo la proprietà (4) richiede una dimostrazione. Sia $\mathcal{U}$ una partizione di $[n + 1]$ in m classi; distinguiamo due casi: se la partizione contiene l'insieme $\{n + 1\}$ formato dal solo elemento

$n+1$ oppure no. Il numero delle partizioni di $[n+1]$ in m classi che contengono l'insieme $\{n+1\}$ formato dal solo elemento $n+1$ è chiaramente uguale al numero $S(n, m-1)$ delle partizioni di $[n]$ in $m-1$ classi. D'altra parte per ogni partizione di $[n]$ in m classi si possono costruire m partizioni di $[n+1]$ in m classi aggiungendo l'elemento $n+1$ ad una delle classi della partizione negli m modi possibili. Dunque il numero delle partizioni di $[n+1]$ in m classi in modo che l'elemento $[n+1]$ non formi da solo una delle classi è $mS(n, m)$. ■

Mostriamo ora un altro esempio di definizione induttiva. Si definisce *l'n-esimo numero di Bell* B_n come il numero di tutte le partizioni dell'insieme $[n]$; dunque

$$\boxed{B_n = \textstyle\sum_{i=0}^{n} S(n,i)}.$$

Usando la formula esplicita per $S(n, i)$ si può ottenere quella per B_n; tuttavia si può dimostrare direttamente con un argomento di carattere combinatorio la seguente formula induttiva per B_n (si osservi che una dimostrazione diretta mediante la definizione esplicita sarebbe estremamente più complicata):

Lemma 3.4.2 *I numeri di Bell B_n ammettono le seguente definizione induttiva: $B_0 = 1$ e*

$$B_{n+1} = \sum_{i=0}^{n} \binom{n}{i} B_i.$$

DIMOSTRAZIONE. Sia U un sottoinsieme $U \subseteq [n+1]$ di $[n+1]$ di cardinalità i e tale che $(n+1) \notin U$. Se $\mathcal{U}$ è una partizione di U, possiamo definire una partizione $\mathcal{U}'$ di $[n+1]$ aggiungendo ad $\mathcal{U}$ il complemento U^c di U, cioè $\mathcal{U}' = \mathcal{U} \cup \{\mathcal{U}^\rfloor\}$. Dunque il numero delle partizioni di $[n+1]$ costruite in tal modo a partire dal sottoinsieme U è B_i. Poichè il numero dei sottoinsiemi $U \subseteq [n+1]$ di cardinalità i e tali che $(n+1) \notin U$ è $\binom{n}{i}$, per $0 \leq i \leq n$, si ha che in tal modo si possono costruire

$$\sum_{i=0}^{n} \binom{n}{i} B_i$$

partizioni di $[n+1]$. D'altra parte *ogni* partizione di $[n+1]$ è in modo unico di questa forma: se $\mathcal{U}'$ è una partizione di $[n+1]$, sia V l'unico elemento di $\mathcal{U}'$ tale che $(n+1) \in V$ e sia U il complemento $U = V^c$ di V; chiaramente $(n+1) \notin U$ e $\mathcal{U}'$ induce una partizione $\mathcal{U}$ di U tale che $\mathcal{U}' = \mathcal{U} \cup \{\mathcal{U}^\sqcup\}$. ■

3.5 Esercizi

1. Sia X un insieme finito e sia $\mathcal{U} = \{\mathcal{U}_\rangle\}_{\rangle \in \mathcal{I}}$ X una famiglia di sottoinsiemi non vuoti e disgiunti U_i di X. Si provi che $\mathcal{U}$ è una partizione di X se e solo se
$$|X| = \sum_{i \in I} |U_i| .$$

2. Si provi che i numeri $S(i,k)$ (per $i,k = 1,\ldots,m$) costituiscono una matrice quadrata invertibile. Indicando gli elementi della matrice inversa con $s(i,k)$, si provi che essi sono numeri interi che soddisfano l' identità:
$$m_{(n)} = \sum_{k=1}^{n} s(n,k) m^k.$$
(Suggerimento: si osservi che dalla formule precedenti si ottiene $m^n = \sum_{k=1}^{m} S(n,k) m_{(k)}$; facendo variare n tra 1 e m, si ha un sistema di m equazioni in m incognite $x_n = m_{(n)}$; i numeri $s(i,k)$ sono chiamati *numeri di Stirling di prima specie*).

3. Indicando con $B(x)$ la funzione generatrice dei numeri di Bell, si provi che dalla loro definizione ricorsiva si ha che $B(x)$ soddisfa l'equazione differenziale $B'(x) = B(x)\exp(x)$ e quindi che, tenendo conto della condizione iniziale $B(0) = 1$, la funzione $B(x)$ è $B(x) = \exp(\exp(x) - 1)$.

3.6 Il Teorema di Lagrange

Un esempio classico di partizione su un insieme è nella teoria dei gruppi. D'ora in poi, quando parliamo genericamente di gruppi, diremo semplicemente "sia G un gruppo ...", indicando semplicemente il nome G

dell'insieme su cui è data l'operazione binaria che soddisfa le proprietà espresse per una operazione di gruppo; non daremo neppure un simbolo particolare per l'operazione, che denoteremo con la semplice giustapposizione; chiameremo "prodotto" l'operazione generica e denoteremo con 1 la sua identità. Ciò premesso, sia G un gruppo e sia $H \subseteq G$ un suo sottogruppo. Diremo "*laterale destro di H in G*" un insieme della forma

$$gH = \{gh \mid h \in H\}$$

cioè un insieme costituito da tutti i prodotti di un elemento $g \in G$ fissato per tutti gli elementi $h \in H$. Denoteremo con G/H l'insieme di tutti i laterali destri.

Teorema 3.6.1 *L'insieme G/H di tutti i laterali destri di H in G è una partizione di G.*

DIMOSTRAZIONE. Cominciamo con l'osservare che due laterali g_1H e g_2H sono uguali (*come sottoinsiemi di G!*) se e solo se $g_2^{-1}g_1 \in H$. Infatti, se $g_1H = g_2H$, poichè $g_1 \in g_1H$ dato che $1 \in H$, allora $g_1 \in g_2H$; cioè esiste $h \in H$ tale che $g_1 = g_2h$ e dunque $g_2^{-1}g_1 = h \in H$. Viceversa, sia $g_2^{-1}g_1 = h \in H$ (e quindi anche $g_1^{-1}g_2 \in H$, perchè H è sottogruppo). Allora se g_1h è un elemento del laterale g_1H, si ha: $g_1h = (g_2g_2^{-1})g_1h = g_2(g_2^{-1}g_1)h$ e l'elemento $(g_2^{-1}g_1)h$ è in H perchè H è chiuso rispetto al prodotto; dunque $g_1H \subseteq g_2H$. Similmente si prova $g_2H \subseteq g_1H$ e dunque che $g_1H = g_2H$.

È ora immediato provare che l'insieme G/H di tutti i laterali destri di H in G è una partizione di G: ogni elemento g di G appartiene al laterale gH, come già osservato. Inoltre, se appartenesse anche ad altro laterale g_1H, cioè se esistesse h tale che $g = g_1h$, allora $g_1^{-1}g = h \in H$ e dunque per quanto dimostrato prima si avrebbe $gH = g_1H$. Dunque ogni elemento di G appartiene ad uno ed un solo laterale. ■

Il Teorema di Lagrange è una applicazione immediata del precedente teorema al caso in cui il gruppo G sia *finito*. Premettiamo che per un gruppo finito, la cardinalità dei suoi elementi è spesso chiamata anche "*ordine*" del gruppo.

Teorema 3.6.2 *(di Lagrange) Se G è un gruppo finito e H è un suo sottogruppo, allora l'ordine di H divide l'ordine di G.*

DIMOSTRAZIONE. La dimostrazione è immediata per il precedente teorema. Per l'esercizio 1 di 3.5, si ha che la cardinalità di G è uguale alla somma delle cardinalità dei laterali distinti, perchè essi costituiscono una partizione di G. Ma ogni laterale gH ha la stessa cardinalità di H, perchè la moltiplicazione per g definisce una funzione

$$H \longrightarrow gH$$

che è ovviamente suriettiva e che è anche iniettiva, dato che in un gruppo l'esistenza dell'inverso garantisce la proprietà di cancellazione. Dunque la cardinaltà di G è la somma della cardinalità di H tante volte quanti sono i laterali distinti, cioè

$$|G| = |G/H||H| \,. \blacksquare$$

Più avanti daremo un esempio che il teorema di Lagrange non può essere invertito: non è detto che per ogni divisore dell'ordine di un gruppo finito esista un sottogruppo di ordine il divisore.

Una immediata conseguenza del teorema di Lagrange è l'osservazione che ogni elemento g di un gruppo finito G elevato all'ordine del gruppo è l'elemento neutro. Infatti, basta osservare che le potenze di g costituiscono un *sottogruppo* di G (perchè?) e dunque che il loro numero è un divisore dell'ordine di G. Questo fatto è alla base di un famoso teorema, il "*piccolo teorema di Fermat*":

Teorema 3.6.3 *Siano a e n due naturali positivi coprimi con $a < n$. Allora $a^{\phi(n)}$ è congruo a 1 modulo n, essendo ϕ la funzione di Eulero definita in 2.9. In particolare, per ogni primo p e per ogni naturale non nullo $a < p$, si ha $a^{p-1} = hp + 1$.*

DIMOSTRAZIONE. Infatti, il fatto che a e n siano coprimi significa che a è invertibile in $\mathbf{Z}_n$; dunque a è un elemento del gruppo degli invertibili di $\mathbf{Z}_n$ e dunque elevato all'ordine di tale gruppo, che è $\phi(n)$, è l'elemento neutro 1, ciò che significa che è congruo a 1 modulo n. $\blacksquare$

3.7 Esercizi

1. Si dimostri che i gruppi finiti privi di sottogruppi propri sono (a meno di isomorfismi) tutti e soli i gruppi $\mathbf{Z}_p$ con p primo.

3.8 Relazioni

Una *relazione* (*binaria*) R tra un insieme X ed un insieme Y è un sottoinsieme

$$R \subseteq X \times Y$$

del prodotto cartesiano $X \times Y$. Useremo anche la notazione

$$xRy$$

o anche

$$R(x, y)$$

(che leggeremo x *è in relazione* R *con* y) per indicare che $\langle x, y\rangle \in R$.

Se $R \subseteq X \times Y$ è una relazione tra X e Y, chiameremo *matrice* di R la sua funzione caratteristica (si veda il paragrafo 1.6)

$$c_R \colon X \times Y \longrightarrow \mathbf{2}.$$

La matrice di R è dunque una funzione che associa ad ogni coppia $\langle x, y\rangle \in X \times Y$ il valore di verità 0 se x non è in relazione R con y ed il valore di verità 1 se x è in relazione R con y. Se X e Y sono gli insiemi finiti standard $[n]$ ed $[m]$, allora la matrice di una relazione R tra X e Y può essere rappresentata con una matrice ad m righe ed n colonne in cui l'elemento di posto (i, k) è 0 o 1 secondo se il k-esimo elemento di $[n]$ è o non è in relazione R con l'i-esimo elemento di $[m]$. Vediamo alcuni esempi:

1. Sia $f \colon X \longrightarrow Y$ una funzione; la relazione

   $$\Gamma(f) = \{\langle x, f(x)\rangle \mid x \in X\} \subseteq X \times Y$$

   è detta *grafo* della funzione f. Si osservi che $\Gamma(f)$ altro non è che l'immagine della funzione $\langle 1, f\rangle : X \longrightarrow X \times Y$. In particolare, il grafo della funzione identità 1_X è la relazione

   $$\{\langle x, x\rangle \mid x \in X\};$$

   indicheremo tale relazione con I_X e la chiameremo *relazione identica su* X; se X è l'insieme finito $[n]$, la matrice di $I_{[n]}$ è la matrice quadrata i cui elementi sulla diagonale principale sono tutti

uguali a 1, mentre tutti gli altri elementi sono 0. In generale, se $f\colon [n] \longrightarrow [m]$ è una funzione tra insiemi finiti standard, allora la matrice del suo grafo ha la proprietà che ogni colonna ha uno ed un solo elemento diverso da 0.

2. Sia $f\colon X \longrightarrow Y$ una funzione; la relazione

$$N(f) = \{\langle x_1, x_2\rangle \in X \times X |\ f(x_1) = f(x_2)\} \subseteq X \times X$$

è detta *nucleo di equivalenza di* f. Ricordando cosa vuol dire che una funzione è iniettiva, si vede che la condizione di iniettività di f equivale a $N(f) = I_X$.

3. La relazione $D \subseteq \mathbf{N} \times \mathbf{N}$ definita da

$$D = \{\langle n, m\rangle \in \mathbf{N} \times \mathbf{N} |\ \text{esiste } k \in \mathbf{N}\colon kn = m\} \subseteq \mathbf{N} \times \mathbf{N}$$

è detta relazione di divisibilità. Se $\langle n, m\rangle \in D$, diremo che n *divide* m e scriveremo $n|m$.

4. Se $(\mathbf{P}X, \cup, \cap)$ è il reticolo dei sottoinsiemi di un insieme X, la relazione

$$\subseteq_X = \{\langle U, V\rangle \in \mathbf{P}X \times \mathbf{P}X |\ U \subseteq V\} \subseteq \mathbf{P}X \times \mathbf{P}X$$

è la *relazione d'ordine* del reticolo. Si osservi che la relazione $\subseteq_X$ è uguale alla relazione

$$\{\langle U, V\rangle \in \mathbf{P}X \times \mathbf{P}X |\ U \cap V = U\}$$

ed anche alla relazione

$$\{\langle U, V\rangle \in \mathbf{P}X \times \mathbf{P}X |\ U \cup V = V\}\,.$$

5. Un'altra relazione che abbiamo già incontrato è la relazione d'ordine su $\mathbf{N}$;

$$\leq = \{\langle i, j\rangle \in \mathbf{N} \times \mathbf{N} |\ i \leq j\} \subseteq \mathbf{N} \times \mathbf{N}.$$

6. Una relazione $R \subseteq X \times X$ è spesso chiamata un *grafo* (*orientato*) su X. Secondo questa terminologia gli elementi di X sono detti *vertici* del grafo, le coppie $\langle x, y\rangle \in R$ sono detti *lati* e, se X è l'insieme finito standard $[n]$, la matrice di R è chiamata *matrice di incidenza* del grafo. Spesso in tal caso può essere utile rappresentare il grafo con una figura, disegnando gli elementi di X come punti del piano e tracciando una freccia da un vertice x ad un vertice y se xRy.

Per le relazioni $R \subseteq X \times Y$ tra due insiemi fissati X e Y si possono fare le solite operazioni di intersezione, unione e complemento di sottoinsiemi. Ma per le relazioni esistono altre due operazioni fondamentali, per descrivere le quali è piu conveniente usare la notazione

$$R\colon X \nrightarrow Y$$

per indicare una relazione $R \subseteq X \times Y$; diremo anche che X è il *dominio* della relazione R e che Y è il suo *codominio*:

1. per ogni relazione $R\colon X \nrightarrow Y$ diciamo relazione *opposta* la relazione $R^\circ\colon Y \nrightarrow X$ definita da:

$$R^\circ = \{\langle y, x\rangle \mid \langle x, y\rangle \in R\} \subseteq Y \times X;$$

2. se $R\colon X \nrightarrow Y$ e $S\colon Y \nrightarrow Z$ sono due relazioni tali che il codominio di R è uguale al dominio di S, definiamo la *composizione* di R e S come la relazione $SR\colon X \nrightarrow Z$ data da:

$$SR = \{\langle x, z\rangle \in X \times Z \mid \text{esiste } y \in Y \text{ tale che } xRy \text{ e } ySz\}.$$

Elenchiamo alcune proprietà di queste operazioni che può essere utile dimostrare come esercizio, sempre usando il principio di estensionalità per i sottoinsiemi:

1. (*associatività della composizione di relazioni*): siano

$$R\colon X_1 \nrightarrow X_2 \quad , \quad S\colon X_2 \nrightarrow X_3 \quad , \quad T\colon X_3 \nrightarrow X_4$$

tre relazioni componibili; allora:

$$(TS)R = T(SR);$$

2. (*identità*): per ogni relazione $R\colon X \not\longrightarrow Y$:

$$RI_X = R = I_Y R;$$

3. (*distributività*):

$$R(S \cup T) = RS \cup RT$$

e se $0_{Y,Z}$ denota la relazione vuota tra Y e Z, ecc., allora

$$0_{Y,Z} R = 0_{X,Z}.$$

4. (*involuzione*):

$$(R^\circ)^\circ = R,\ (SR)^\circ = R^\circ S^\circ,\ (R \cap S)^\circ = R^\circ \cap S^\circ,\ (R^c)^\circ = (R^\circ)^c;$$

5. (*modularità*):

$$SR \cap T \subseteq (S \cap TR^\circ)R.$$

Un teorema (dovuto a P. Freyd) garantisce che tali proprietà costituiscono una *caratterizzazione elementare* delle relazioni, nel senso *ogni* proprietà elementare delle relazioni è conseguenza di questi cinque assiomi sulle operazioni precedentemente descritte.

Nel caso delle relazioni tra gli insiemi finiti standard $[n]$, le operazioni "opposta" di una relazione e "composizione di relazioni" si interpretano sulle matrici corrispondenti, nelle operazioni di "trasposta" di una matrice e di "prodotto" di matrici. Tuttavia, trattandosi di matrici di valori di verità, l'usuale prodotto di matrici (righe per colonne) è definito mediante le *operazioni logiche* sui valori di verità e non su quelle numeriche; dunque, il "prodotto" di valori di verità, che è la congiunzione logica $\wedge$, coincide con il prodotto numerico sugli interi 0 e 1, ma la "somma" di valori di verità, che è la disgiunzione logica $\vee$, *non* coincide completamente con la somma numerica, perchè $1 \vee 1 = 1$.

Dunque, se $A = (a_{i,k})$ e $B = (b_{k,l})$ sono due matrici di valori di verità di tipo (m, n) e (n, p) rispettivamente, l'elemento $c_{i,l}$ di posto (i, l) nella matrice prodotto AB è dato da:

$$c_{i,l} = \bigvee_{k=1}^{n} (a_{i,k} \wedge b_{k,l}).$$

Il linguaggio delle relazioni permette di esprimere molti fatti sulle funzioni. È senz'altro un utile esercizio dimostrare i seguenti fatti:

1. Sia $f: X \longrightarrow Y$ una funzione e sia $\Gamma(f): X \longrightarrow Y$ il suo grafo; allora:
$$\Gamma(f)\Gamma(f)^\circ \subseteq I_Y \quad , \quad \Gamma(f)^\circ\Gamma(f) \supseteq I_X.$$
Viceversa, una relazione $R: X \longrightarrow Y$ è il grafo di un 'unica funzione $X \longrightarrow Y$ se $RR^\circ \subseteq I_Y$ e $R^\circ R \supseteq I_X$.

2. La relazione $N(f)$ nucleo di equivalenza di una funzione f può essere espressa come:
$$N(f) = \Gamma(f)^\circ\Gamma(f);$$
Dunque f è iniettiva se e solo se
$$\Gamma(f)^\circ\Gamma(f) = I_X.$$

3. Una funzione f è suriettiva se e solo se
$$\Gamma(f)\Gamma(f)^\circ = I_Y.$$

3.9 Esercizi

1. Un grafo orientato G su un insieme X si dice *riflessivo* quando contiene tutte le coppie $\langle x, x\rangle$ per $x \in X$ e si dice *irriflessivo* quando non ne contiene nessuna. In altre parole, $G \subseteq X \times X$ è riflessivo quando $I_X \subseteq G$ ed è irriflessivo quando $I_X \cap G = \emptyset$. Se $|X| = [n]$, si conti il numero dei grafi orientati riflessivi e quello dei grafi orientati irriflessivi.

2. Sia $R \subseteq X \times X$ una endorelazione di X. R è detta *simmetrica* se $R = R^\circ$. In particolare, se $X = [n]$, R è simmetrica se e solo se la sua matrice è una matrice simmetrica. Si provi che il numero delle endorelazioni simmetriche su $[n]$ è
$$2^{\frac{n(n+1)}{2}}.$$

3. Un *grafo* su un insieme X è un sottoinsieme delle coppie *non ordinate* di X che sia irriflessivo. Si conti il numero dei grafi su un insieme X di cardinalità n; (Suggerimento: un grafo su X è semplicemente un sottoinsieme dell'insieme dei sottoinsiemi di X con *due* elementi distinti, dunque ...). Si provi che i grafi su X sono in corrispondenza biunivoca con le endorelazioni simmetriche di X che non intersecano I_X.

4. Se X e Y sono monoidi (gruppi, anelli) e se $f: X \longrightarrow Y$ è un omomorfismo di monoidi (di gruppi, di anelli), si provi che $N(f) \subseteq X \times X$ è un sottomonoide (sottogruppo, sottoanello) del monoide (gruppo, anello) prodotto.

3.10 Relazioni di Equivalenza

Una importante nozione è quella di "*relazione di equivalenza*" su un insieme X. L'idea è quella che in molte circostanze si è condotti naturalmente a definire su un insieme X una endorelazione $E \subseteq X \times X$ che ha le stesse proprietà formali della relazione di uguaglianza (se $\langle x_1, x_2 \rangle \in E$, scriveremo $x_1 \sim_E x_2$ e diremo che x_1 è E-equivalente a x_2):

1. (*riflessività*): per ogni $x \in X$: $x \sim_E x$.

2. (*simmetria*): se $x_1 \sim_E x_2$, allora $x_2 \sim_E x_1$.

3. (*transitività*): se $x_1 \sim_E x_2$ e $x_2 \sim_E x_3$, allora $x_1 \sim_E x_3$.

Un esempio immediato è il nucleo di equivalenza

$$N(f) = \{\langle x_1, x_2 \rangle \ \in X \times X \mid f(x_1) = f(x_2)\} \subseteq X \times X$$

di una funzione $f: X \longrightarrow Y$. Possiamo dire che il dato della funzione f ci conduce a considerare una diversa nozione di uguaglianza sugli elementi di X: due elementi x_1 e x_2 vengono considerati uguali, non più quando sono lo stesso elemento di X, ma quando *diventano* lo stesso elemento di Y dopo l'applicazione della funzione f.

Un altro importante esempio è il seguente: se H è un sottogruppo di un gruppo G, definiamo la relazione

$$\sim_H \subseteq G \times G$$

(che chiameremo "*congruenza modulo* H"), nel modo seguente:

$$g_1 \sim_H g_2 \text{ se e solo se } g_1^{-1} g_2 \in H.$$

È facile verificare che la congruenza modulo H è una relazione di equivalenza su G *le cui classi di equivalenza sono proprio i laterali destri di* H *in* G. Infatti, $x \in [g]_{\sim_H}$ se e solo se $x^{-1}g \in H$, se e solo se $g^{-1}x \in H$, se e solo se esiste $h \in H$ tale che $g^{-1}x = h$, se e solo se esiste $h \in H$ tale che $x = gh$, se e solo se $x \in gH$.

Gli esempi sono in realtà moltissimi. Per convincerci, cerchiamo di contare quante sono le relazioni di equivalenza su un insieme finito. Come al solito, un conto diretto è assai difficile, ma un argomento basato sull'idea che una relazione di equivalenza su un insieme X è in un certo senso come introdurre una nuova relazione di uguaglianza tra gli elementi di X, ci fornisce immediatamente la risposta. Infatti, se $E \subseteq X \times X$ è una relazione di equivalenza su X e se $x \in X$ è un elemento di X, allora, pensando che E è una nuova "uguaglianza" su X, dobbiamo convenire che tutti gli elementi x' tali che $x \sim_E x'$ devono essere identificati con x e che dunque il sottoinsieme

$$[x]_E = \{x' \in X \mid x_\sim x'\} \subseteq X$$

di X va considerato come un unico elemento di un insieme su cui E diventa la relazione di uguaglianza abituale, nel senso che per quanto riguarda la nuova "uguaglianza" E, essi non sono distinguibili. L'insieme $[x]_E$ viene detto "*classe di equivalenza di* x" ed il fatto cruciale è il seguente:

Teorema 3.10.1 *Se E è una relazione di equivalenza su X, allora l'insieme*

$$X/E$$

delle classi di equivalenza è una partizione di X, che viene detto "insieme quoziente di X rispetto ad E".

DIMOSTRAZIONE. Per la proprietà riflessiva di E, ogni elemento $x \in X$ appartiene a $[x]_E$. Inoltre, se $x \in [y]_E$, allora $[x]_E = [y]_E$. Infatti, se $z \in [x]_E$, cioè se $x \sim_E z$, allora poichè $x \in [y]_E$, cioè $y \sim_E x$, per la transitività si ha $y \sim_E z$, cioè $z \in [y]_E$; dunque $[x]_E \subseteq [y]_E$; similmente, usando anche la simmetria si dimostra che $[y]_E \subseteq [x]_E$: se $z \in [y]_E$, cioè se $y \sim_E z$, allora poichè $y \sim_E x$ e dunque per la simmetria anche $x \sim_E y$, per la transitività si ha $x \sim_E z$, cioè $z \in [x]_E$. Dunque $[x]_E = [y]_E$. ■

La costruzione X/E delle classi di equivalenza di E definisce una funzione dall'insieme delle relazioni di equivalenza su X all'insieme delle partizioni di X. Il fatto essenziale è che tale funzione è *biunivoca*: se $\mathcal{U} = (X_i)_{i \in I}$ è una partizione di X, allora la relazione $\mathcal{U}$ definita da

$$x_1 \sim_{\mathcal{U}} x_2 \text{ se e solo se esiste } i \in I \text{ tale che } x_1 \in X_i \text{ e } x_2 \in X_i$$

è una relazione di equivalenza (esercizio). È facile ora mostrare che la funzione così definita dalle partizioni di X alle relazioni di equivalenza su X è la funzione inversa della precedente e che dunque tali funzioni stabiliscono *una corrispondenza biunivoca tra le relazioni di equivalenza su X e le partizioni di X*. In particolare, se $X = [n]$, allora il numero delle relazioni di equivalenza su $[n]$ è il numero di Bell B_n.

La precedente discussione implica anche che l'esempio di relazione di equivalenza dato dal nucleo di equivalenza di una funzione è in realtà *l'unico esempio*, nel senso che data una qualsiasi relazione di equivalenza E su un insieme X, esiste sempre una funzione il cui nucleo di equivalenza è la relazione E. Infatti, se E è una relazione di equivalenza su X, allora possiamo costruire l'insieme quoziente X/E e possiamo considerare la funzione

$$p_E \colon X \longrightarrow X/E$$

definita da

$$p_E(x) = [x]_E \,.$$

Tale funzione è detta "*proiezione sul quoziente*" ed è facile vedere che *il suo nucleo di equivalenza è proprio E*:
$N(p_E) = \{\langle x, y\rangle \mid p_E(x) = p_E(y)\} = \{\langle x, y\rangle \mid [x]_E = [y]_E\} =$
$= \{\langle x, y\rangle \mid x \sim_E y\} = E,$

poichè, per quanto dimostrato nel precedente teorema, $[x]_E = [y]_E$ se e solo se $x \sim_E y$. Tutto ciò conduce a dimostrare la seguente *proprietà universale della costruzione dell'insieme quoziente*:

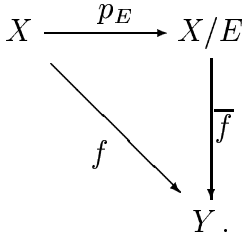
Teorema 3.10.2 *Se E è una relazione di equivalenza su X, allora la proiezione*

$$p_E \colon X \longrightarrow X/E$$

di X sull'insieme quoziente X/E ha la seguente proprietà (universale): per ogni funzione $f \colon X \longrightarrow Y$ tale che $E \subseteq N(f)$, cioè tale che se $x \sim_E y$ allora $f(x) = f(y)$, esiste un'unica funzione $\overline{f} \colon X/E \longrightarrow Y$ tale che

$$\overline{f}(p_E(x)) = f(x)\,,$$

ciò che può essere espresso dicendo che il seguente diagramma commuta:



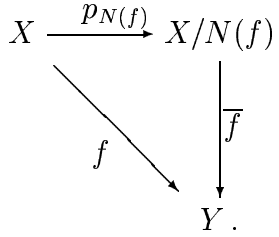
DIMOSTRAZIONE. Se vogliamo che la funzione $\overline{f}$ sia definita in modo che il diagramma commuti, l'unica possibilità è quella di definirla su ogni classe di equivalenza $[x]_E$ come:

$$\overline{f}([x]_E) = f(x)\,.$$

non rimane dunque che mostrare che la precedente definizione è *ben posta*, cioè che *non dipende dal rappresentante della classe di equivalenza*: se y è un altro rappresentante della classe $[x]_E$, cioè se $[x]_E = [y]_E$, allora $f(x) = f(y)$. Ricordando che $[x]_E = [y]_E$ se e solo se $x \sim_E y$, questa è proprio la condizione espressa nell'enunciato. Questa condizione si esprime anche dicendo che "*f è costante sulle classi di equivalenza*" o anche che "*f è una funzione E-invariante*". ■

Come esempio di applicazione di questo teorema, possiamo considerare una qualunque funzione $f \colon X \longrightarrow Y$ ed il suo nucleo di equivalenza

$E = N(f)$. Poichè $E = N(f)$, per il teorema precedente abbiamo una fattorizzazione di f



Per come è definita, la funzione $\overline{f}$ si fattorizza attraverso l'immagine $\mathrm{Im}(f) \subseteq Y$ di f

$$\overline{f}: X/N(f) \longrightarrow \mathrm{Im}(f)$$

e si dimostra che è un *isomorfismo*: è suriettiva, perchè dato $y \in \mathrm{Im}(f)$, cioè dato un $y \in Y$ della forma $f(x)$ per un qualche $x \in X$, allora $\overline{f}([x]_{N(f)}) = f(x) = y$; è iniettiva, perchè se $\overline{f}([x]_{N(f)}) = \overline{f}([y]_{N(f)})$, allora $f(x) = f(y)$, dunque $x \sim_{N(f)} y$, ciò che equivale a $[x]_{N(f)} = [y]_{N(f)}$. Abbiamo così dimostrato il

Teorema 3.10.3 *Ogni funzione $f: X \longrightarrow Y$ induce un isomorfismo*

$$\overline{f}: X/N(f) \longrightarrow Im(f)$$

tra l'insieme quoziente del dominio con il nucleo di equivalenze $N(f)$ di f e l'immagine $Im(f)$ di f.

3.11 Esercizi

1. Se X è un insieme e $R \subseteq X \times X$ è una relazione su X, si provi che

   a) la relazione
   $$R_s = R \cup R^{\circ}$$
   è una relazione simmetrica ed è la più piccola relazione simmetrica che contiene R;

   b) la relazione
   $$R_r = I_X \cup R$$
   è una relazione riflessiva ed è la più piccola relazione riflessiva che contiene R;

c) se R è una relazione riflessiva e simmetrica, la relazione

$$E(R) = \bigcup_n R^n \,,$$

dove R^n indica la *composizione di R iterata n volte*, è la più piccola relazione di equivalenza che contiene R.

2. Se $[4]$ denota l'insieme standard con 4 elementi e se R è la relazione binaria su $[4]$ rappresentata dalla matrice

$$\begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

si calcoli la matrice che rappresenta la più piccola relazione di equivalenza che contiene R.

3. Se M è un monoide commutativo con cancellazione, si consideri la relazione $\sim$ su $X = M \times M$ definita da:

$$\langle x_1, y_1 \rangle \sim \langle x_2, y_2 \rangle \text{ se e solo se } x_1 y_2 = x_2 y_1 \,.$$

Si provi che

a) la relazione $\sim$ è una relazione di equivalenza su X;

b) l'operazione definita sulle classi di equivalenza da

$$[x, y]_\sim [u, v]_\sim = [xu, yv]_\sim$$

non dipende dalla scelta dei rappresentanti e dunque definisce una operazione binaria su $X/\sim$, che è associativa, commutativa e dotata di un elemento neutro;

c) il monoide commutativo $X/\sim$ è in realtà un *gruppo* e la funzione

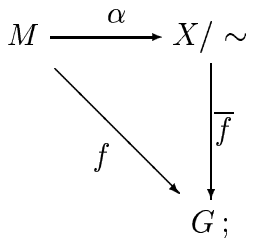
$$\alpha \colon M \longrightarrow X/\sim$$

definita da

$$\alpha(m) = [\langle m, 1 \rangle]_\sim$$

è un omomorfismo *iniettivo*.

d) vale la seguente proprietà universale: se $f: M \longrightarrow G$ è un omomorfismo dal monoide M ad un gruppo commutativo G, allora esiste un'unico omomorfismo di gruppi $\overline{f}: X/\sim \longrightarrow G$ tale che il seguente diagramma commuti:



Per questa ragione il gruppo $X/\sim$ è detto "*gruppo libero sul monoide commutativo cancellativo* M".

d) se M è il monoide commutativo $(\mathbf{N}, +)$ dei naturali rispetto alla somma, allora $X/\sim$ è isomorfo all gruppo $(\mathbf{Z}, +)$ dei numeri interi e, se M è il monoide $(\mathbf{N}^{>0}, \cdot)$ dei naturali non nulli rispetto al prodotto, allora $X/\sim$ è isomorfo al gruppo moltiplicativo dei razionali positivi.

3.12 Congruenze e Strutture Quoziente

Poniamoci ora la questione di capire cosa succede quando applichiamo la precedente costruzione dell'insieme quoziente ad un insieme su cui è assegnata una struttura algebrica di monoide, di gruppo o altra ancora. Più precisamente, limitandoci al caso dei monoidi per fissare le idee, il problema è il seguente: supponiamo che $E \subseteq M \times M$ sia una relazione di equivalenza e che su M sia assegnata una struttura di monoide $(M, \cdot, 1)$; il problema è:

determinare le condizioni necessarie e sufficienti sulla relazione E *affinchè sull'insieme quoziente* M/E *sia definibile una struttura di monoide in modo che la proiezione* $p_E: M \longrightarrow M/E$ *sia un omomorfismo.*

Ricordando che la proiezione p_E è definita da $p_E(x) = [x]_E$ e che $N(p_E) = E$, la condizione necessaria è evidente: se su M/E è definibile una struttura di monoide in modo che p_E sia un omomorfismo, allora

$E = N(p_E) \subseteq M \times M$ è un *sottomonoide del monoide prodotto* $M \times M$ (si veda l'esercizio 4 di 3.9). Viceversa, se $E \subseteq M \times M$ è una relazione di equivalenza su M che inoltre sia un sottomonoide del monoide prodotto $M \times M$, allora dato che questa condizione significa in particolare che

$$x \sim_E x' \quad \text{e} \quad y \sim_E y' \rightarrow xy \sim_E x'y'$$

si può definire una operazione sulle classi di equivalenza come

$$[x]_E[y]_E = [xy]_E \,.$$

Infatti, la condizione per E di essere sottomonoide del prodotto *equivale* alla condizione che la precedente definizione dell'operazione sulle classi di equivalenza sia *ben posta*, cioè non dipenda dalla scelta dei rappresentanti:

se $[x]_E = [x']_E$ *e* $[y]_E = [y']$, *allora* $[xy]_E = [x'y']_E$,

poichè due classi di equivalenza sono uguali se e solo se i rappresentanti sono equivalenti. È ora immediato (esercizio) dimostrare che tale operazione è associativa, che ha un elemento neutro (la classe di equivalenza $[1]_E$ dell'elemento neutro) e che la proiezione $p_(x) = [x]_E$ è un omomorfismo. Osserviamo infine che se $E \subseteq M \times M$ è una relazione riflessiva, allora la condizione di chiusura rispetto alla operazione binaria è sufficiente a garantire che E è un sottomonoide, poichè la riflessività assicura che l'elemento neutro $\langle 1, 1\rangle$ è in E. In definitiva, possiamo riassumere e completare la precedente discussione nel seguente teorema:

Teorema 3.12.1 i) *Dato un monoide M ed una relazione di equivalenza E su M, condizione necessaria e sufficiente sulla relazione E affinchè sull'insieme quoziente M/E sia definibile una struttura di monoide in modo che la proiezione $p_E\colon M \longrightarrow M/E$ sia un omomorfismo è che E sia un sottomonoide del monoide prodotto $M \times M$, ciò che equivale a:*

$$x \sim_E x' \quad e \quad y \sim_E y' \rightarrow xy \sim_E x'y' \,.$$

*Diremo in tal caso che la relazione di equivalenza è una "*congruenza*" sul monoide M.*

ii) *Se E è una congruenza sul monoide M, allora la proiezione*

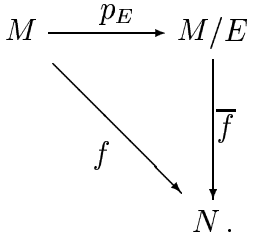
$$p_E\colon M \longrightarrow M/E$$

di M sul monoide quoziente M/E ha la seguente proprietà (universale):

per ogni omomorfismo di monoidi $f\colon M \longrightarrow N$ tale che $E \subseteq N(f)$, cioè tale che se $x \sim_E y$ allora $f(x) = f(y)$, esiste un unico omomorfismo $\overline{f}\colon M/E \longrightarrow N$ tale che

$$\overline{f}(p_E(x)) = f(x)\,,$$

ciò che può essere espresso dicendo che il seguente diagramma commuta:



Dimostrazione. Solo la seconda parte necessita di una dimostrazione, poichè la prima parte è stata dimostrata nella discussione precedente il teorema. A questo scopo, basta dimostrare che nelle condizioni espresse dal teorema l'unica funzione $\overline{f}\colon M/E \longrightarrow N$ tale che $\overline{f}(p_E(x)) = f(x)$, è in realtà un omomorfismo, cioè che

$$\overline{f}([x]_E[y]_E) = \overline{f}([x]_E)\overline{f}([y]_E) \quad , \quad \overline{f}([1]_E) = 1\,.$$

Ricordando la definizione di $\overline{f}$ data dal teorema precedente

$$\overline{f}([x]_E) = f(x)$$

e che

$$[x]_E[y]_E = [xy]_E\,,$$

perchè E è una congruenza, per la definizione di $\overline{f}$ tali condizioni sono proprio le condizioni che f sia un omomorfismo di monoidi. ■

3.13 Esercizi

1. Sia M un semigruppo commutativo e sia $R \subseteq M \times M$ la relazione binaria su M definita da

$$xRy =_{df} \text{ esiste } t \in M \text{ tale che } xt = yt\,.$$

Si provi che

a) la relazione R è una congruenza di semigruppi.

b) il semigruppo quoziente M/R è un semigruppo in cui vale la proprietà di cancellazione.

c) La proiezione $p\colon M \longrightarrow M/R$ soddisfa la seguente proprietà universale: se $f\colon M \longrightarrow S$ è un omomorfismo da M verso un semigruppo in cui vale la proprietà di cancellazione, esiste un unico omomorfismo $\tilde{f}\colon M/R \longrightarrow S$ tale che $\tilde{f}p = f$. Per questa ragione il semigruppo M/R viene detto "*semigruppo cancellativo libero su M*".

d) Se M è un monoide commutativo, anche M/R è un monoide commutativo e la proiezione p è un omomorfismo di monoidi.

2. Sia M un monoide commutativo con cancellazione e si consideri la relazione $\sim \subseteq X \times X$ sul prodotto $X = M \times M$ definita nell'esercizio 3.11.3. Si provi che

a) Se $\sim$ è una congruenza sul monoide prodotto $X = M \times M$.

b) Se M è un monoide commutativo, non necessariamente cancellativo, si possono eseguire in questo ordine le due costruzioni del monoide cancellativo libero su M e al risultato applicare la costruzione del gruppo libero sul monoide commutativo cancellativo M/R. Il risultato è un gruppo commutativo, denotato con $k(M)$ e detto "*gruppo di Grothendieck di M*". Ricordando la proprietà universale descritta nell'esercizio 3.11.3 d), si enunci la proprietà universale di $k(M)$.

3.14 Sottogruppi normali

Se $E \subseteq G \times G$ è una congruenza su un monoide G e se G è in particolare un gruppo, allora affinchè il monoide quoziente G/E sia un gruppo (cioè affinchè E sia una congruenza di gruppi) c'è da aspettarsi che E debba possedere ulteriori proprietà. Infatti, se G/E è un gruppo, allora dato che la proiezione $p_E: G \longrightarrow G/E$ è un omomorfismo di gruppi, allora $E = N(p_E)$ è un sottogruppo del gruppo prodotto $G \times G$ e dunque in particolare è chiuso rispetto all'inverso:

$$x \sim_E y \to x^{-1} \sim_E y^{-1}\,.$$

Tuttavia, è facile vedere che se E è una congruenza di monoide e se G è un gruppo, allora E è una congruenza di gruppo:
se $x \sim_E y$, allora poichè per la riflessività si ha $y^{-1} \sim_E y^{-1}$, segue che $y^{-1}x \sim_E y^{-1}y = 1$ e dunque, ancora per la riflessività applicata a x^{-1}, che $y^{-1} \sim_E x^{-1}$; infine, per la simmetria, si ha $x^{-1} \sim_E y^{-1}$.

Dunque, la nozione di congruenza per monoidi e per gruppi coincide e, sostituendo uniformemente "gruppo" a "monoide" e "omomorfismo di gruppi" a quello di "omomorfismo di monoide", il teorema del paragrafo precedente può essere dimostrato anche per i gruppi. C'è tuttavia una differenza cruciale tra la nozione di congruenza per i monoidi e quella per i gruppi: il fatto che *per i gruppi la nozione di congruenza è rappresentabile*, nel senso seguente:

Teorema 3.14.1 *Dato un gruppo G, il prendere il nucleo $ker(p_E)$ della proiezione sul quoziente di G per un congruenza E su G stabilisce una corrispondenza biunivoca tra l'insieme delle congruenze su G e l'insieme dei sottogruppi "normali" di G, cioè di quei sottogruppi N di G con la proprietà*

$$n \in N, g \in G \Rightarrow g^{-1}ng \in N\,.$$

L'elemento $g^{-1}ng$ si chiama "coniugato di n mediante g" e dunque la proprietà di un sottogruppo di essere normale si esprime a parole dicendo che il coniugato di ogni elemento di N mediante ogni elemento di G appartiene a N. Se N è un sottogruppo normale di G, scriveremo $N \lhd G$.

DIMOSTRAZIONE. Osserviamo subito che il nucleo di un *qualunque* omomorfismo di gruppi $f: G \longrightarrow H$ è un sottogruppo normale: infatti, se $n \in \ker(f)$, cioè se $f(n) = 1$, allora qualunque sia $g \in G$, si ha $f(g^{-1}ng) = f(g^{-1})f(n)f(g) = f(g^{-1})f(g) = f(1) = 1$, cioè $g^{-1}ng \in \ker(f)$. Dunque prendere il nucleo della proiezione sul quoziente di G per una congruenza associa ad ogni congruenza su G un sottogruppo normale di G.
Viceversa, abbiamo già visto in un precedente esercizio che se N è un sottogruppo di un gruppo G, la relazione

$$\sim_N \subseteq G \times G$$

(che abbiamo chiamato "*congruenza modulo* N"), definita da

$$g_1 \sim_N g_2 \text{ se e solo se } g_1^{-1}g_2 \in N,$$

è una relazione di equivalenza su G *le cui classi di equivalenza sono proprio i laterali destri di* N *in* G. Vediamo ora che tale relazione di equivalenza è una congruenza se e solo se N è un sottogruppo normale di G. Infatti, se la relazione di congruenza modulo N è una congruenza, allora la proiezione sul quoziente è un omomorfismo di gruppi e dunque il suo nucleo è un sottogruppo normale; ma $\ker(p_{\sim_N})$ è proprio N. Viceversa, sia N un sottogruppo normale; mostriamo che la relazione di congruenza modulo N è una congruenza: se $x_1 \sim_N x_2$ e $y_1 \sim_N y_2$, cioè se $x_1^{-1}x_2 \in N$ e $y_1^{-1}y_2 \in N$, allora $y_1^{-1}x_1^{-1}x_2y_2 \in N$, poichè $y_1^{-1}x_1^{-1}x_2y_2 = [y_1^{-1}(x_1^{-1}x_2)y_1]y_1^{-1}y_2$; ma l'espressione dentro la parentesi quadra è un elemento di N perchè N è normale e per una delle due ipotesi, mentre l'altro fattore è in N per l'altra delle due ipotesi e, dato che N è un sottogruppo, il loro prodotto è in N. Dunque $y_1^{-1}x_1^{-1}x_2y_2 \in N$, ciò che precisamente significa $x_1y_1 \sim_N x_2y_2$, cioè che la relazione di equivalenza "congruenza modulo N" è una congruenza. ■.

È chiaro che per i gruppi abeliani ogni sottogruppo è normale, ma si puó mostrare che in generale non è vero per i gruppi non commutativi. Ad esempio, calcoliamo tutti i sottogruppi del gruppo Δ_4 delle simmetrie del quadrato e poi ricerchiamo quali sono quelli normali. Per semplificare questi calcoli, vedremo che sarà utile una caratterizzazione dei sottogruppi normali in termini di laterali:

Teorema 3.14.2 *Un sottogruppo N di un gruppo G è normale se e solo se ogni laterale destro è uguale al corrispondente laterale sinistro: per ogni $g \in G$, $gN = Ng$.*

DIMOSTRAZIONE. È chiaro che se per ogni $g \in G$, $gN = Ng$, allora per ogni $g \in G$ si ha $g^{-1}Ng = N$ e dunque per ogni $n \in N$, $g^{-1}ng \in N$, cioè N è normale.

Viceversa, se N è normale, e se $gn \in gN$, allora $gn = (gng^{-1})g$; ma $gng^{-1} \in N$, perchè è il coniugato di n mediante g^{-1} e N è normale; dunque gn è della forma mg, con $m = gng^{-1} \in N$, dunque $gn \in Ng$; perciò $gN \subseteq Ng$. Similmente si prova $Ng \subseteq gN$ e dunque $gN = Ng$. ■. A titolo di esemplificazione cerchiamo ora di determinare tutti i sottogruppi e tutti i sottogruppi normali di un gruppo particolare, Δ_4. Cominciamo con l'osservare che per il teorema di Lagrange, i sottogruppi possono avere come ordine solo un divisore dell'ordine di Δ_4, che è 8, dunque quelli propri solo 2 e 4. I sottogruppi di ordine 2 sono della forma $H = \{1, x\}$, dove x è un elemento per cui $x^2 = 1$. Ora, in Δ_4 tali elementi x sono solo R^2, D, RD, R^2D e R^3D.

Per quanto riguarda i sottogruppi di ordine 4, il sottoinsieme delle rotazioni $T = \{1, R, R^2, R^3\}$ è evidentemente un sottogruppo di ordine 4 isomorfo al gruppo $\mathbf{Z}_4$. Per trovare gli altri eventuali sottogruppi di ordine 4, ragioniamo così: se X è un sottogruppo di ordine 4, allora $X \cap T$ è un sottogruppo di X e di T, dunque può avere come ordine solo 1, 2 o 4, sempre per il teorema di Lagrange. Non può essere 1, perchè allora X dovrebbe contenere tre elementi di Δ_4 che non sono in T, cioè tre fra i quattro elementi $\{D, RD, R^2D, R^3D\}$; una semplice verifica mostra che ogni possibile scelta (sono 4) non è chiusa rispetto al prodotto. Non può essere 4, perchè altrimenti $X = T$. Dunque può essere solo 2. Poichè $X \cap T$ è un sottogruppo di ordine 2 di T, può essere solo $\{1, R^2\}$; dunque X deve avere la forma $X = \{1, R^2, x, y\}$, dove x e y sono due qualsiasi elementi di Δ_4. Ma possiamo subito escludere che x o y siano R o R^3, perchè in tal caso per la chiusura di X rispetto al prodotto si avrebbe $X = T$. Dunque x e y possono essere solo uno dei rimanenti 4 elementi.

Se $x = D$, deve per forza essere $y = R^2D$ (perchè X deve essere chiuso rispetto al prodotto) ed è facile verificare che il sottoinsieme

$$U = \{1, R^2, D, R^2D\}$$

è chiuso rispetto al prodotto ed è perciò un sottogruppo di Δ_4.

Se $x = RD$, allora y deve essere per forza R^3D e si può verificare che il sottoinsieme

$$V = \{1, R^2, RD, R^3D\}$$

è chiuso rispetto al prodotto, dunque è un sottogruppo di Δ_4.

Se $x = R^2D$, allora y deve essere D e si ricade perciò nel caso di U.

Infine, se $x = R^3D$, allora y deve essere RD e si ricade ancora nel caso di U.

Dunque non ci possono essere altri sottogruppi di Δ_4 di ordine 4 oltre a quelli trovati: T, U e V.

Riassumendo, Δ_4 ha 8 sottogruppi non banali, 3 di ordine 4 e 5 di ordine 2. Resta da decidere quali di questi sottogruppi sono normali. Iniziamo da quelli di ordine 4. Un criterio per decidere se un sottogruppo è normale è quello di verificare se i laterali sinistri coincidono con i laterali destri. Per il teorema di Lagrange, poichè l'ordine del gruppo è 8, i laterali di un sottogruppo di ordine 4 sono solo due: il sottogruppo stesso ed il laterale che si ottiene da un qualsiasi altro elemento che non appartenga al sottogruppo. Ma tale laterale deve per forza coincidere con il complemento insiemistico del sottogruppo e dunque laterali destri e sinistri coincidono. Perciò ogni sottogruppo di ordine 4 è normale. Si osservi che tale ragionamento si può generalizzare a Δ_n, per n qualsiasi e quindi si può dimostrare che il sottogruppo delle rotazioni di un qualsiasi gruppo diedrico è normale.

Consideriamo ora i sottogruppi di ordine 2. Cominciamo con il sottogruppo $X = \{1, R^2\}$. I quattro laterali sinistri di X sono:
$X1 = X; XR = \{R, R^3\}; XD = \{D, R^2D\}; X(RD) = \{RD, R^3D\}$;
i corrispondenti laterali destri sono:
$X; RX = XR; DX = \{D, DR^2\} = \{D, R^2D\} = XD; (RD)X = \{RD, RDR^2\} = \{RD, R^3D\} = X(RD)$;
dunque i laterali sinistri e destri coincidono, perciò X è normale.

Vediamo che X è l'unico sottogruppo normale di ordine 2. Infatti, il coniugato di D mediante R è $RDR^{-1} = RDR^3 = RDRR^2 = RR^3DR^2 = DR^2 = R^2D$; dunque il sottogruppo $\{1, D\}$ non è normale perchè non contiene il coniugato di D mediante R. Similmente, il coniugato di RD mediante $D = D^{-1}$ è $DRDD = DR = R^3D$; dunque il sottogruppo $\{1, RD\}$ non è normale perchè non contiene il coniugato

di RD mediante D. In modo simile si verifica che anche gli altri due sottogruppi di ordine 2 non sono normali.

Concludendo, degli otto sottogruppi propri di Δ_4, solo quattro sono normali: i tre sottogruppi di ordine 4 e il sottogruppo $X = \{1, R^2\}$ di ordine 2.

Rimangono da determinare i gruppi quoziente di Δ_4. Per quanto riguarda i sottogruppi normali di ordine 4, poichè i laterali sono solo 2 ed esiste una sola struttura di gruppo su un insieme con due elementi, i gruppi quoziente i gruppi quoziente sono tutti isomorfi a $\mathbf{Z}_2$; è chiaro però che le tre proiezioni $\Delta_4 \longrightarrow \mathbf{Z}_2$ sono tutte diverse.

Consideriamo l'unico sottogruppo normale $X = \{1, R^2\}$ di ordine 2. Gli elementi del gruppo quoziente Δ_4/X sono i laterali $X_0 = X$, $X_1 = XR$, $X_2 = XD$ e $X_3 = X(RD)$. Calcoliamo la tavola di moltiplicazione di Δ_4/X:

	X_0	X_1	X_2	X_3
X_0	X_0	X_1	X_2	X_3
X_1	X_1	X_0	X_3	X_2
X_2	X_2	X_3	X_0	X_1
X_3	X_3	X_2	X_1	X_0

Ad esempio: $X_0X_0 = (XR)(XR) = XR^2 = X = X_0$, perchè $R^2 \in X$; $X_3X_1 = (X(RD))(XR) = X(RDR) = XD = X_2$, ecc.

Possiamo identificare Δ_4/X ad un gruppo noto. Consideriamo il gruppo prodotto $\mathbf{Z}_2 \times \mathbf{Z}_2$; i suoi elementi sono $0 = \langle 0, 0\rangle$, $a = \langle 1, 0\rangle$, $b = \langle 0, 1\rangle$ e $c = \langle 1, 1\rangle$; la tavola di moltiplicazione è

	0	a	b	c
0	0	a	b	c
a	a	0	c	b
b	b	c	0	a
c	c	b	a	0

Si vede immediatamente che ponendo $X_0 = 0$, $X_1 = a$, $X_2 = b$ e $X_3 = c$, le tavole di moltiplicazione di Δ_4/X e di $\mathbf{Z}_2 \times \mathbf{Z}_2$ coincidono; dunque la funzione $f(X_0) = 0$, $f(X_1) = a$, $f(X_2) = b$, $f(X_3) = c$ è un isomorfismo di gruppi:

$$f\colon \Delta_4/X \longrightarrow \mathbf{Z}_2 \times \mathbf{Z}_2\,.$$

Il gruppo $\mathbf{Z}_2 \times \mathbf{Z}_2$ è detto "gruppo trirettangolo" e può essere interpretato geometricamente come il gruppo delle simmetrie del rettangolo, cioè l'identità, il ribaltamento intorno all'asse a, quello intorno all'asse b e la loro composizione (che è la rotazione di 180 gradi). Se chiamiamo rispettivamente 0, a, b e c queste quattro simmetrie e scriviamo la loro tavola di moltiplicazione, troviamo esattamente la tavola di moltiplicazione di $\mathbf{Z}_2 \times \mathbf{Z}_2$.

Si osservi che conoscendo il significato geometrico del gruppo prodotto $\mathbf{Z}_2 \times \mathbf{Z}_2$ avremmo potuto aspettarci il fatto che Δ_4/X è isomorfo a $\mathbf{Z}_2 \times \mathbf{Z}_2$. Infatti, la relazione di equivalenza associata al sottogruppo $\{1, R^2\}$ consiste nel dichiarare equivalenti due simmetrie del quadrato quando o sono uguali o si ottengono una dall'altra attraverso una rotazione di 180 gradi. Ma esaminando le simmetrie del quadrato si può vedere che ognuna di esse è una simmetria del quadrato considerato come rettangolo, oppure, se non lo è, lo è invece quella che si ottiene dopo una rotazione di 180 gradi.

Per finire, si osservi che, sebbene il gruppo Δ_4 non sia commutativo, lo sono invece tutti i suoi quozienti propri.

3.15 Esercizi

1. Si consideri il gruppo $\mathcal{Q}$ generato da due elementi i e j soddisfacenti le relazioni:

$$i^4 = 1 \ , \ i^2 = j^2 \ , \ ji = i^3 j \, .$$

   Si dimostri che $\mathcal{Q}$ ha 8 elementi e si scriva la tavola di moltiplicazione. Si determinino tutti i suoi sottogruppi e si dica quali di essi sono normali. Il gruppo $\mathcal{Q}$ è detto "*gruppo dei quaternioni*".

2. Se $N_1 \lhd G_1$ e $N_2 \lhd G_2$ sono due sottogruppi normali di due gruppi G_1 e G_2, si provi che il prodotto $N_1 \times N_2$ è un sottogruppo normale del prodotto $G_1 \times G_2$ e che esiste un isomorfismo canonico

$$(G_1 \times G_2)/(N_1 \times N_2) \sim G_1/N_1 \times G_2/N_2 \, .$$

3.16 Ideali

Rimane da capire che cosa è una congruenza per gli anelli. Poichè un anello è un gruppo abeliano $(A,+)$ dotato di un prodotto $(A,\cdot)$ distributivo rispetto all'operazione di "somma" del gruppo abeliano, una *congruenza di anello* è una congruenza di gruppo abeliano, quindi un sottogruppo $I \subseteq A$ del gruppo $(A,+)$, tale che la congruenza che genera è anche una congruenza per il monoide $(A,\cdot)$.

Teorema 3.16.1 *Dato un anello $(A,+,\cdot)$, un sottogruppo $I \subseteq A$ del gruppo $(A,+)$ è tale che la congruenza che genera è anche una congruenza per il monoide $(A,\cdot)$ se e solo se soddisfa la seguente condizione:*
per ogni $a \in A$ e per ogni $i \in I$, $ia \in I$ e $ai \in I$,
ciò che si esprime dicendo che I è un "ideale (bilatero)".

Dimostrazione. Se la congruenza generata da I è anche una congruenza rispetto alla struttura di monoide, allora il gruppo quoziente A/I è anche un anello rispetto al *prodotto* di laterali

$$(I+a)(I+b) =_{df} I+ab$$

e la proiezione $p_I\colon A \longrightarrow A/I$ è un omomorfismo di anelli il cui nucleo è proprio il sottogruppo I. Ma è un fatto generale che se $f\colon A \longrightarrow B$ è un omomorfismo di anelli, allora $\ker(f)$ è un ideale bilatero: $\mathrm{Ker}(f)$ è un sottogruppo di $(A,+)$ e se $i \in \ker(f)$ e $a \in A$, allora $f(ai) = f(a)f(i) = f(a)0 = 0$, cioè $ai \in \ker(f)$ e similmente $ia \in \ker(f)$.

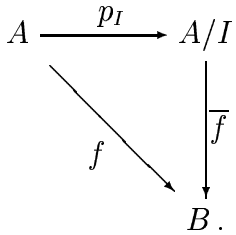
Viceversa, sia I un ideale dell'anello A e mostriamo che la congruenza generata da I è una congruenza anche rispetto al monoide $(A,\cdot)$. Ricordiamo che la congruenza generata da I è definita da
$a \sim b$ se e solo se $(b-a) \in I$.
Il fatto che tale relazione di equivalenza sia anche una congruenza rispetto al prodotto significa che
se $a \sim b$ e $c \sim d$, allora $ac \sim bd$,
cioè
se $(b-a) \in I$ e $(d-c) \in I$, allora $(bd-ac) \in I$.
Quest'ultima proprietà si dimostra nel modo seguente:
$bd-ac = bd+ad-adac = (b-a)d + a(d-c)$,

che è un elemento di I, poichè $(b-a) \in I$, $(d-c) \in I$ e poichè I essendo un ideale bilatero è chiuso rispetto al prodotto a sinistra e a destra con ogni elemento di A ed anche rispetto alla somma. ■

Dunque, se I è un ideale bilatero di un anello A, il gruppo quoziente A/I è in realtà un anello e la proiezione $p_I: A \longrightarrow A/I$ è un omomorfismo di anelli, che soddisfa la proprietà universale:
per ogni anello B ed ogni omomorfismo di anelli $f: A \longrightarrow B$ *tale che* $I \subseteq ker(f)$, esiste un unico omomorfismo $\overline{f}: A/I \longrightarrow B$ tale che

$$\overline{f}(p_I(x)) = f(x)\,,$$

ciò che può essere espresso dicendo che il seguente diagramma commuta:



In seguito ci occuperemo prevalentemente di anelli *commutativi*; in tal caso la nozione di ideale si semplifica poichè basta richiedere solo una delle due condizioni di chiusura rispetto al prodotto per gli elementi dell'anello. Moltissimi sono gli esempi di ideali: nel precedente teorema si prova che il nucleo di ogni omomorfismo di anelli è un ideale; di più, si prova che infatti *ogni* ideale appare in tal modo, ad esempio come nucleo dell'omomorfismo dato dalla proiezione sul quoziente. Di particolare interesse sono gli ideali definiti nel modo seguente: sia A un anello commutativo e sia $a \in A$; si vede facilmente (esercizio) che l'insieme

$$(a) = \{xa \mid x \in A\}$$

è un ideale di A, chiamato "*ideale principale generato da* a".

Capitolo 4

Azioni

4.1 M-insiemi

Se $\mathrm{M} = \mathrm{End}(X)$ è il monoide delle endofunzioni di un insieme X, c'è un tipo di *operazione* evidente di M su X

$$* : \mathrm{End}(X) \times X \longrightarrow X\,,$$

la valutazione

$$f * x = f(x)\,,$$

che evidentemente soddisfa le proprietà

M_1) $g * (f * x) = (gf) * x$

M_2) $1_X * x = x$.

In generale chiameremo *azione (sinistra)* di un monoide M su un insieme X ogni operazione ("*azione*") di tipo $* : M \times X \longrightarrow X$, il cui valore su una coppia $\langle m, x\rangle$ denoteremo con $m * x$, che soddisfa M_1 e M_2. Diremo che "X è un M-insieme (sinistro)" o anche che "M opera (a sinistra) su X". Gli esempi sono moltissimi. Eccone solo alcuni:

1. Prendendo per X il monoide stesso M, l'operazione di M è una azione $M \times M \longrightarrow M$: le identità M_1 e M_2 sono semplicemente l'associatività e l'unità sinistra.

2. Se N è un sottomonoide di M e $* : M \times X \longrightarrow X$ è una azione di M su X, allora si ha per restrizione una azione $N \times X \longrightarrow X$. Ci sono due esempi particolari di tale situazione:

   a) Se N è un sottomonoide di un monoide M e $M \times M \longrightarrow M$ è l'azione dell'esempio 1), allora essa induce per restrizione una azione
   $$N \times M \longrightarrow M$$
   di N su M; quando M è un gruppo e N è un suo sottogruppo, questa azione è già stata considerata nel Teorema di Lagrange (si veda 3.6).

   b) se $\mathrm{Aut}(X)$è il sottogruppo di $\mathrm{End}(X)$ degli automorfismi di un insieme X, allora per restrizione dell'azione considerata all'inizio del paragrafo si ottiene una azione
   $$\mathrm{Aut}(X) \times X \longrightarrow X\,.$$

3. Se X è un insieme e $X^n = X \times \cdots \times X$ è l'insieme delle n-uple ordinate degli elementi di X, allora c'è una azione del gruppo $S([n]) = S_n$ delle permutazioni su $[n]$ su X^n
   $$* : S_n \times X^n \longrightarrow X^n$$
   definita da
   $$\sigma * \langle x_1, \ldots, x_n \rangle = \langle x_{\sigma(1)}, \ldots, x_{\sigma(n)} \rangle\,.$$

4. Se G è un gruppo, si può definire una azione
   $$* : G \times G \longrightarrow G$$
   mediante la formula
   $$g * x = gxg^{-1}\,.$$
   Tale azione è detta '*coniugio*' e si riduce alla proiezione
   $$G \times G \longrightarrow G$$
   se e solo se il gruppo G è abeliano.

5. Il fatto che il monoide che agisce su un insieme X sia un gruppo ha diverse conseguenze non banali. Ad esempio, una è la seguente: se un monoide M agisce su un insieme X mediante una azione $*: M \times X \longrightarrow X$, allora agisce anche sull'insieme $\mathbf{P}X$ delle parti di X mediante la formula

$$m * U = \{m * x \,|\, x \in U\} \subseteq X\,,$$

per ogni sottoinsieme U di X. Ora, se M è un gruppo, allora per ogni elemento m la funzione $m * (-): X \longrightarrow X$ è un isomorfismo, poichè la funzione inversa è la funzione $m^{-1} * (-)$, ciò che si dimostra dagli assiomi M_1 e M_2; dunque la funzione $m* -$ induce un isomorfismo da U a $m * U$. Pertanto, denotando con $\mathbf{P}_n X$ l'insieme dei sottoinsiemi di X di cardinalità n, quando M è un gruppo, l'azione sulle parti di X si restringe ad una azione sulle parti di cardinalità n.

6. Molti sono gli esempi di natura geometrica e qui ci limitiamo a citarne due. Se $\mathcal{P}$ denota l'insieme dei punti del piano e $\mathbb{T}$ il gruppo delle traslazioni, c'è una azione evidente

$$+: \mathbb{T} \times \mathcal{P} \longrightarrow \mathcal{P}$$

che ha la proprietà: per ogni coppia P_1 e P_2 di punti esiste un'unica traslazione T, tale che $T + P_1 = P_2$, ciò che autorizza a denotare tale T con $T = P_2 - P_1$. A sua volta, il gruppo delle dilatazioni $\mathbb{D}$ agisce in modo naturale sul gruppo delle traslazioni (come?).

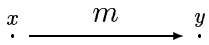
7. (Macchine di Turing) Se $f: X \longrightarrow X$ è una funzione, c'è una azione

$$\mathbf{N} \times X \longrightarrow X$$

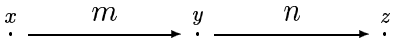
del monoide $(\mathbf{N}, +, 0)$ definita da $n * x = f^n(x)$. Inoltre, *ogni azione* $*: N \times X \longrightarrow X$ *del monoide dei naturali rispetto alla somma è di questo tipo per un'unica funzione* $f: X \longrightarrow X$ *(esercizio; si definisca* $f(x) = 1 * x$*...).*

È spesso assai utile associare ad una azione un grafo che la arricchisce di una intuizione geometrica, il suo *diagramma*. Se $*: M \times X \longrightarrow X$

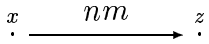
è una azione, consideriamo il seguente grafo Diag$(*: M \times X \longrightarrow X)$: i *vertici* sono gli elementi di $x, y, z, ...$ di X, che denoteremo con punti; i *lati* (che sono orientati e che dunque denoteremo con frecce)



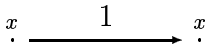
sono gli elementi m tali che $m * x = y$. Si osservi che tale grafo ha la seguente ulteriore struttura: se



sono due lati come descritti nella figura, allora mn è un lato



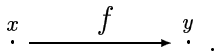
che chiameremo *composizione* dei due lati e che per ogni elemento x l'unità 1 del monoide M è un lato



Il lettore può facilmente provare che gli assiomi di azione equivalgono al fatto che la composizione è *associativa*, quando è definita, e che ha *identità*. Diverse significative proprietà di una azione possono essere espresse mediante proprietà ('geometriche') del suo diagramma. Ad esempio, il fatto che il diagramma sia un grafo *connesso*, cioè che ogni coppia di vertici sia congiunta da almeno un lato, equivale alla proprietà

dell'azione di essere *transitiva*, cioè che per ogni coppia di elementi $x, y \in X$, esiste $m \in M$ tale che $m * x = y$.

La costruzione del diagramma di una azione ha una conseguenza assai utile. Cominciamo con il precisare che cosa è un *grafo (orientato)*, ciò che fino ad ora non abbiamo fatto, avendo preferito lasciare all'intuizione del lettore immaginare la sua corretta definizione. Un grafo orientato è il dato di due insiemi, l'insieme V dei vertici e quello L dei lati, e di due funzioni $d_0, d_1 \colon L \longrightarrow V$ che assegnano ad ogni lato f due vertici $d_0(f)$ e $d_1(f)$, detti rispettivamente dominio e codominio di f. Per visualizzare un grafo, disegneremo i suoi vertici come punti e, se f è un lato di cui $x = d_0(f)$ e $y = d_1(f)$ sono il dominio ed il codominio, disegneremo f come freccia



Ad ogni grafo orientato $\mathbb{G} = (d_0, d_1 \colon L \longrightarrow V)$ possiamo sempre associare una relazione binaria $R(\mathbb{G}) \subseteq V \times V$ sull'insieme V dei suoi vertici semplicemente prendendo l'immagine della funzione

$$\langle d_0, d_1 \rangle \colon L \longrightarrow V \times V \, .$$

Dunque due vertici sono in relazione se e solo se esiste un lato che li congiunge. Un fatto assai utile è il seguente:

Teorema 4.1.1 *Se $\mathbb{G}$ è il grafo dato dal diagramma dell'azione $* \colon G \times X \longrightarrow X$ di un gruppo G su un insieme X, allora la relazione $R(\mathbb{G}) \subseteq X \times X$ è una relazione di equivalenza. Denoteremo con X/G l'insieme quoziente, con una notazione che sottintende l'azione di G su X.*

La semplice dimostrazione è lasciata al lettore. Qui ci limitiamo a qualche esempio. Gli elementi dell'insieme quoziente X/G sono le classi di equivalenza della relazione di equivalenza $R(\mathbb{G}) \subseteq X \times X$ e dunque sono gli insiemi

$$G * x = \{ g * x \mid g \in G \} \, ,$$

che vengono detti *"orbite di x"*. In particolare si osservi che nell'esempio della restrizione ad un sottogruppo H di G dell'azione data dall'operazione di G, le orbite sono proprio i laterali di H in G. In un certo senso, l'insieme X/G è una "misura" dell'azione: una prima rozza approssimazione del senso di tale affermazione è il seguente, di facile verifica: una azione è transitiva se e solo se X/G ha un solo elemento, cioè se e solo se esiste una sola orbita. Il lettore provi a mostrare che l'esempio 2, b) è un esempio di azione transitiva.
Concludiamo convenendo che se M^{op} denota il *monoide opposto* di M, cioè il monoide sullo stesso insieme M, ma in cui l'operazione è definita da $x \cdot y =_{def} yx$, allora una azione sinistra di M^{op} su un insieme X è detta *azione destra* di M su X. Dunque una azione destra di M su X è una funzione $* : M \times X \longrightarrow X$ tale che

M_1^{op}) $m * (n * x) = (nm) * x$

M_2) $1 * x = x$.

4.2 Rappresentazioni, centro

Un monoide *"concreto"* è un monoide di endofunzioni, cioè un monoide della forma $\mathrm{End}(X)$ per un insieme X. Una *rappresentazione* di un monoide M è un qualsiasi modo di compararlo con un monoide concreto, dunque semplicemente un *omomorfismo di monoidi*

$$\mathrm{M} \xrightarrow{\quad\rho\quad} \mathrm{End}(X)\,,$$

per un qualche insieme X. Quando ρ è *iniettiva*, si dice che la rappresentazione è *fedele* e in tal caso si può pensare ad una rappresentazione come ad un "modello" di M. È semplice ma assai utile osservare che il principio di λ-conversione (si veda 1.12) permette di dimostrare il seguente

Teorema 4.2.1 *Per ogni monoide* M *ed ogni insieme X, la λ-conversione induce una corrispondenza biunivoca tra le rappresentazioni*

$$\mathrm{M} \xrightarrow{\quad\rho\quad} \mathrm{End}(X)\,,$$

e le azioni

$$* : M \times X \longrightarrow X\,.$$

DIMOSTRAZIONE. Basta mostrare che data una rappresentazione ρ, la sua trasposta $* = \rho^t : M \times X \longrightarrow X$ soddisfa gli assiomi di azione:

$(mn) * x = \rho^t(mn, x) = \rho(mn, x) = \rho[m, \rho(n, x)] = \rho^t[m, \rho^t(n, x] =$
$= m * (n * x)$,

e che data una azione $* : M \times X \longrightarrow X$, la sua aggiunta esponenziale $\lambda* : M \longrightarrow \mathrm{End}(X)$ è un omomorfismo, cioè una rappresentazione di M:

$\lambda * (mn)(x) = (mn) * x = m * (n * x) = \lambda * (m)[(\lambda * (n)(x)] =$
$= [\lambda * (m)][(\lambda * (n)](x)$. ■

Osserviamo che se M è in particolare un gruppo G e ρ è una rappresentazione, allora ogni valore di ρ è un *automorfismo* di X e dunque una rappresentazione di un gruppo G è in realtà un *omomorfismo di gruppi*

$$\mathrm{G} \xrightarrow{\ \rho\ } \mathrm{Aut}(X)$$

di G al gruppo degli automorfismi di un insieme X.

Il caso speciale della aggiunta esponenziale dell'azione data dalla operazione stessa $M \times M \longrightarrow M$ è una rappresentazione canonica detta *rappresentazione di Cayley* di M. Se M è in particolare un gruppo G, allora la sua rappresentazione di Cayley

$$\mathrm{G} \xrightarrow{\ c\ } \mathrm{Aut}(G)$$

è sempre *fedele* (esercizio).

In generale, non è detto che una rappresentazione sia fedele. Ad esempio, consideriamo la rappresentazione

$$\mathrm{G} \xrightarrow{\ \rho\ } \mathrm{Aut}(G)$$

aggiunta esponenziale dell'azione di coniugio $\rho(g)(x) = gxg^{-1}$. Il suo nucleo, che sappiamo essere un sottogruppo normale di G, è:

$\ker(\rho) = \{g \in G \mid \rho(g) = 1_G\} = \{g \in G \mid \rho(g)(x) = x,\ \text{per ogni } x \in G\} =$

$= \{g \in G \mid gxg^{-1} = x, \text{ per ogni } x \in G\} =$
$= \{g \in G \mid gx = xg, \text{ per ogni } x \in G\} = Z(\mathrm{G})$.

Il sottogruppo normale $Z(\mathrm{G})$ di G è detto "*centro*" di G e coincide con G se e solo se G è abeliano. Possiamo considerare il gruppo quoziente

$$\mathrm{G}/Z(\mathrm{G})$$

che per il teorema di isomorfismo è un gruppo isomorfo all'immagine $\mathrm{Im}(\rho)$ e che pertanto ha una rappresentazione fedele nel gruppo $\mathrm{Aut}(G)$. Il sottogruppo $\mathrm{Im}(\rho)$ di $\mathrm{Aut}(G)$ è detto "*sottogruppo degli automorfismi interni di* G", nome che trova la sua spiegazione in una altra importante proprietà della rappresentazione di coniugio: per ogni elemento $g \in G$, l'isomorfismo $\rho(g)$ è un *isomorfismo di gruppi*, dunque un automorfismo di G *come gruppo* (esercizio). Una conseguenza importante della precedente osservazione è la seguente. Poichè per ogni $g \in G$, l'isomorfismo $\rho(g)\colon G \longrightarrow G$ è un isomorfismo di gruppi, dunque porta sottogruppi in sottogruppi, l'azione di coniugio si estende all'insieme $Sub(\mathrm{G})$ dei sottogruppi di G

$$*\colon G \times Sub(\mathrm{G}) \longrightarrow Sub(\mathrm{G})$$

definendo $g * H = gHg^{-1}$ ed è un semplice esercizio provare direttamente gli assiomi di azione.

La precedente osservazione sulla rappresentazione data dal coniugio è di carattere generale: data *ogni* rappresentazione

$$\mathrm{G} \xrightarrow{\ \rho\ } \mathrm{Aut}(X)$$

di un gruppo G, possiamo sempre prendere l'immagine $\mathrm{Im}(\rho)$, che è isomorfa al quoziente di G per il nucleo di ρ e ottenere una rappresentazione fedele del quoziente nel gruppo degli automorfismi dello stesso insieme X.

4.3 Stabilizzatori

Se $*\colon M \times X \longrightarrow X$ è una azione di un monoide M su un insieme X, allora per ogni $x \in X$ il sottoinsieme di M

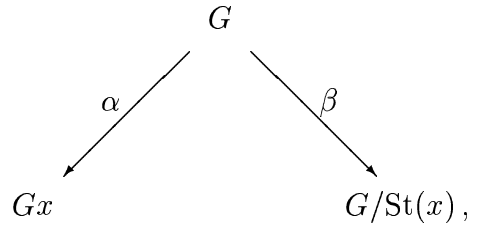
$$\mathrm{St}(x) = \{m \in M \mid m * x = x\}$$

dato dagli elementi la cui azione lascia fisso x è un *sottomonoide* di M e, quando M è un gruppo G, un sottogruppo di G (esercizio) detto "*stabilizzatore di x*". Diremo che x è un *invariante* per l'azione di G, o che è G-*invariante*, se $\mathrm{St}(x) = \mathrm{G}$, cioè se l'azione di ogni elemento di G lascia fisso x. Ad esempio, se consideriamo l'azione di coniugio, lo stabilizzatore $\mathrm{St}(x)$ è il sottogruppo di G

$$\mathrm{St}(x) = \{g \in G \,|\, gxg^{-1} = x\} = \{g \in G \,|\, gx = xg\}\,.$$

Esso viene detto "*centralizzante*" di x e spesso viene denotato con $\mathrm{C}(x)$. Dunque $x \in Z(\mathrm{G})$ se e solo se $\mathrm{C}(x)$=G e perciò gli elementi G-invarianti per l'azione di coniugio sono proprio gli elementi del centro. Quando consideriamo l'azione di coniugio estesa ai sottogruppi, gli elementi G-invarianti sono precisamente i sottogruppi normali.

D'ora in poi per una azione $G \times X \longrightarrow X$ semplificheremo la notazione denotando il suo valore su una coppia $\langle g, x\rangle$ semplicemente con gx. Data dunque una azione di un gruppo G su un insieme X, consideriamo le due funzioni



aventi lo stesso dominio G e aventi codominio rispettivamente l'orbita Gx di x e l'insieme quoziente di G per il sottogruppo stabilizzatore di x, definite da $\alpha(g) = gx$ e $\beta(g) = g\mathrm{St}(x)$. Entrambe sono suriettive e hanno inoltre lo stesso nucleo di equivalenza:

$\alpha(g_1) = \alpha(g_2) \Leftrightarrow g_1x = g_2x \Leftrightarrow g_2^{-1}g_1x = x \Leftrightarrow g_2^{-1}g_1 \in \mathrm{St}(x)$
$\Leftrightarrow g_1\mathrm{St}(x) = g_2\mathrm{St}(x) \Leftrightarrow \beta(g_1) = \beta(g_2)$.

Per la proprietà universale dei quozienti esiste un unico isomorfismo

$$\theta : Gx \longrightarrow G/\mathrm{St}(x)$$

tale che $\theta\alpha = \beta$. In particolare, quando X e G sono *finiti*, passando alle cardinalità si ottiene:

$$|Gx| = |G/\mathrm{St}(x)| = [\mathrm{G{:}St}(x)]\,,$$

da cui ricordando che le orbite sono una partizione di X, si ha *l'equazione delle classi*:

$$|X| = \sum_{x \in X/G} [\mathrm{G{:}St}(x)] = \sum_{x \in X/\mathrm{G}} \frac{|\mathrm{G}|}{|\mathrm{St}(x)|} ,$$

dove la notazione $x \in X/\mathrm{G}$ significa che la somma è estesa ad una scelta qualsiasi di $x \in X$, uno ed uno solo per ogni orbita.

Illustriamo l'equazione delle classi per l'azione di coniugio di in gruppo finito G. Ricordiamo che in tal caso gli stabilizzatori $\mathrm{St}(x)$ sono i centralizzanti $\mathrm{C}(x)$, mentre le orbite, dette "*classi coniugate*" e spesso denotate con $\mathrm{Cl}(x)$, sono i sottoinsiemi $\mathrm{Cl}(x) = \{gxg^{-1} \mid g \in G\}$. L'equazione delle classi in tal caso è:

$$|G| = \sum \frac{|G|}{|\mathrm{C}(x)|} ,$$

dove la somma è estesa ad una scelta arbitraria di elementi $x \in G$, uno ed uno solo per ogni classe coniugata.

Vediamo ora alcune applicazioni della teoria fin qui delineata.

Teorema 4.3.1 *Se* G *è un gruppo finito di ordine p^n per un primo p, allora il centro di* G *contiene un elemento diverso dall'unità.*

Dimostrazione. Per il teorema di Lagrange, l'ordine del centralizzante $\mathrm{C}(x)$ di ogni elemento x di G divide l'ordine di G, che è p^n, dunque è una potenza p^{n_x} di p con $n_x \leq n$. Scriviamo l'equazione delle classi per l'azione di coniugio:

$$p^n = \sum_{n_x \leq n} \frac{p^n}{p^{n_x}} = |Z(\mathrm{G})| + \sum_{n_x < n} \frac{p^n}{p^{n_x}} ,$$

dove la somma è estesa ad una scelta arbitraria di elementi $x \in G$, uno ed uno solo per ogni classe coniugata e si è usato che $x \in Z(\mathrm{G})$ se e solo se $\mathrm{C}(x) = G$, se e solo se $p^{n_x} = p^n$, se e solo se $n_x = n$. Dunque

$$|Z(\mathrm{G})| = p^n - \sum_{n_x < n} p^{n - n_x}$$

e pertanto p divide l'ordine del centro, che perciò non può essere ridotto alla sola unità. ■

Corollario 4.3.1 *Ogni gruppo finito di ordine p^2 per un primo p è abeliano.*

DIMOSTRAZIONE. Per il teorema precedente, il centro non è ridotto alla sola unità, dunque il suo ordine deve essere p o p^2. Se l'ordine è p^2 il teorema è vero. Se l'ordine è p, sia a un elemento che non appartiene al centro e consideriamo il suo centralizzante $\mathrm{C}(a)$. Il centro è certamente un sottogruppo di $\mathrm{C}(x)$ e poichè a è in $\mathrm{C}(a)$ e non nel centro, è un sottogruppo proprio. Ancora, l'ordine di $\mathrm{C}(a)$ può essere solo p o p^2, ma non può essere, perchè in ogni caso contraddiremmo l'ipotesi che a non sia nel centro. Dunque G coincide con il suo centro e perciò è abeliano. ■

4.4 Il gruppo simmetrico S_n

La struttura del gruppo degli automorfismi $\mathrm{Aut}(X)$ di un insieme X è completamente determinata dalla cardinalità di X. Infatti un isomorfismo $\phi\colon X \longrightarrow Y$ induce un isomorfismo di gruppi

$$\Phi\colon \mathrm{Aut}(X) \longrightarrow \mathrm{Aut}(Y)$$

mediante $\Phi(\alpha) = \phi\alpha\phi^{-1}$ (esercizio). In particolare, se X è un insieme finito di cardinaltà n, allora $\mathrm{Aut}(X)$ è un gruppo isomorfo al gruppo $\mathrm{Aut}([n]) = \mathrm{S}_n$, detto anche gruppo *simmetrico* di $[n]$.

Vogliamo ora determinare completamente la struttura di tale gruppo. Cosa questo significhi, risulterà chiaro alla fine della discussione e pertanto non ci dilunghiamo a spiegarlo in anticipo. Dovrà anche risultare chiaro da quanto esporremo la ragione del nome "gruppo simmetrico". Il metodo che useremo sarà essenzialmente basato sulle proprietà della azione canonica

$$\mathrm{S}_n \times [n] \longrightarrow [n]$$

dell'esempio 2, b) di 4.1, e delle sue restrizioni. Cominciamo con il ricordare una precedente osservazione secondo cui tale azione è *transitiva*; infatti, per ogni coppia di elementi $i, j \in [n]$ esiste una permutazione $\sigma_{i,j}$ di S_n che manda i in j, la permutazione che "*scambia*" i con j e lascia fissi tutti gli altri elementi. Vedremo che alcuni di tali automorfismi elementari, che soddisfano a condizioni che mostreremo che li

autorizza a chiamare "*simmetrie*", sono i mattoni con cui può venire completamente descritto ogni automorfismo di $[n]$. Per ora osserviamo che essi soddisfano tutti ad una identità, quella che li qualifica come *involuzioni*:

$$\sigma_{ij}^2 = 1_{[n]} .$$

Gli scambi, o trasposizioni, sono particolari permutazioni, dette "*cicli*". Ricordando che in un gruppo finito G, come S_n, ogni elemento g ha un periodo finito $o(g)$, i cicli sono definiti come segue:

Definizione 4.4.1 *Un "ciclo" è una permutazione $\sigma \in S_n$ tale che*

$$o(g) = |\mathrm{Fix}(\sigma)^c| ,$$

cioè tale che il suo ordine è uguale al numero degli elementi di $[n]$ non lasciati fissi, dunque mossi, da σ. Il valore comune k è chiamato "lunghezza del ciclo".

Cerchiamo ora di descrivere come è fatto un ciclo σ di lunghezza k. Sia i_1 il primo elemento di $[n]$ che è mosso da σ; $\sigma(i_1) = i_2$ è a sua volta mosso da σ, come pure ogni $\sigma^m(i_1) = i_{m+1}$ è mosso da σ, per $m < k-1$, poichè σ è una funzione iniettiva. Quando $m = k-1$, ci sono k elementi mossi da σ e *poichè k è anche l'ordine di σ* non può che aversi $\sigma^k(i_1) = i_1$ e si ricomincia da capo. Dunque, σ opera sul sottoinsieme $\{i_1, i_2, \ldots, i_k\}$ di $[n]$, *in questo ordine*, mandando ciascun elemento nel successivo e *opera sul suo complemento come l'identità*; viceversa, se $\{i_1, i_2, \ldots, i_k\}$ è un sottoinsieme di $[n]$ e lo ordiniamo con l'ordine indotto dall'ordine naturale sugli indici, esiste un solo ciclo di lunghezza k che opera su $[n]$, mandando ogni elemento del sottoinsieme con l'ordine scelto sul successivo e lasciando fissi tutti gli altri elementi che non sono nel sottoinsieme. Di qui il nome di ciclo: un ciclo permuta ciclicamente gli elementi del sottoinsieme dato con l'ordine prescritto, mandando ciascuno nel successivo e l'ultimo nel primo e lascia fissi tutti gli altri elementi non nel sottoinsieme. Dunque il numero dei cicli di lunghezza k è il numero dei sottoinsiemi di cardinalità k per il numero degli *ordini lineari* su un insieme di cardinalità k, che altro non è che il numero $n_{(k)}$ delle funzioni iniettive $[k] \longrightarrow [n]$; tale numero deve essere diviso per k, poichè su un sottoinsieme di $[n]$ di cardinalità k esistono esattamente k ordini lineari che danno luogo allo stesso ciclo.

Per facilitare la comprensione e per eseguire calcoli è certamente utile usare la notazione *a due righe* per le permutazioni

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma_1 & \sigma_2 & \dots & \sigma_n \end{pmatrix}$$

(dove per semplificare la notazione abbiamo consistentemente scritto σ_i a posto di $\sigma(i)$) e ad *ad una riga*

$$\gamma = \begin{pmatrix} i_1 & i_2 & \dots & i_k \end{pmatrix}$$

per indicare un ciclo di lunghezza k con l'ordine prescritto. Abbiamo ora tutti gli elementi per dimostrare un primo teorema di struttura, salvo questa ultima osservazione. Due permutazioni σ e τ si dicono "*disgiunte*" se $\mathrm{Fix}(\sigma)^c \cap \mathrm{Fix}(\tau)^c = \emptyset$, cioè se gli elementi mossi da una non sono mossi dall'altra. È facile (e lo lasciamo al lettore) dimostrare il seguente

Lemma 4.4.1 *Se σ e τ sono disgiunte, allora $\sigma\tau = \tau\sigma$.*

Teorema 4.4.1 *Ogni permutazione $\sigma \in \mathrm{S}_n$ si decompone in modo unico (a meno dell'ordine) nel prodotto*

$$\sigma = \gamma_1\gamma_2 \dots \gamma_h$$

dove le permutazioni γ_i sono cicli disgiunti a due a due.

DIMOSTRAZIONE. Sia (σ) il sottogruppo ciclico di S_n generato da σ e consideriamo la restrizione a (σ) dell'azione canonica su $[n]$. Sappiamo che le orbite

$$(\sigma)i = \{i, \sigma(i), (\sigma^2)(i), \dots, (\sigma^{k-1})(i)\}$$

dove k è l'ordine di σ, sono una *partizione* di $[n]$. Attenzione, la cardinaltà di $(\sigma)i$ può essere inferiore a k, poichè sebbene l'ordine di σ sia k, per qualche *particolare* valore i si può avere $(\sigma^r)(i) = i$, anche per un r minore di k ed in tal caso la cardinalità di $(\sigma)i$ è inferiore a k. Il lettore può usare la formula di 4.3 per dare una formula per la cardinalità di ciascuna orbita (esercizio). A questo punto la parte

riguardante la decomposizione enunciata nel teorema è dimostrata: se ci sono h orbite, denotiamole con $C_1, C_2, \ldots, C_h$, allora ad ogni orbita associamo il ciclo determinato dall'ordine delle potenze di σ e otteniamo così h cicli $\gamma_1, \gamma_2, \ldots, \gamma_h$ che hanno la proprietà enunciata dal teorema: se $j \in [n]$, allora j sta in una e una sola orbita, diciamo C_s, e dunque $(\gamma_1, \gamma_2, \ldots, \gamma_h)(j) = \gamma_s)(j) = \sigma(j)$.

D'altra parte, se $\sigma = \delta_1\delta_2 \ldots \delta_l$ è un'altra decomposizione di σ in cicli a due a due disgiunti, allora gli elementi mossi da ciascun ciclo δ_r formano un'orbita C_r di σ e quindi δ_r è il ciclo γ_r. Dunque le due decomposizioni differiscono solo per l'ordine dei fattori. ■

Osservando che per due elementi permutabili e di periodo finito di un gruppo, l'ordine del prodotto è il è il *minimo comune multiplo* degli ordini dei fattori, si può usare il precedente teorema per calcolare l'ordine di una permutazione qualsiasi: esso infatti risulta essere il minimo comune multiplo degli ordini dei cicli disgiunti in cui può venire decomposto in modo unico, ricordando che per definizione l'ordine di un ciclo è semplicemente il numero degli elementi che sono mossi.

Gli scambi, o trasposizioni, sono dunque cicli di lunghezza due. Di particolare interesse sono gli $n-1$ scambi

$$\tau_i = (i, i+1)\,,$$

che chiameremo "simmetrie", perchè sono un sistema di *generatori*, per cui è possibile trovare un sistema di *relazioni* particolarmente significative che giustificano tale nome, che forniscono una presentazione del gruppo simmetrico S_n.

Teorema 4.4.2 *Ogni permutazione di* S_n *si può scrivere come prodotto di cicli della forma* τ_i *e il prodotto in* S_n *è determinato completamente dalle relazioni*

$$\tau_i\tau_j = \begin{cases} 1 & \text{se } j = i \\ \tau_j\tau_i & \text{se } j < i-1 \\ \tau_j\tau_i\tau_j\tau_i & \text{se } j = i-1\,. \end{cases}$$

DIMOSTRAZIONE. È immediato dimostrare (esercizio per il lettore) che tali relazioni valgono per le simmetrie τ in S_n.

Per mostrare che le simmetrie e le relazioni date descrivono completamente S_n come gruppo, cominciamo con l'osservare che *ogni* scambio

(i,j), con $j > i+1$, può essere ottenuto come prodotto dei τ. Si ha infatti $\tau_{i+1}\tau_i\tau_{i+1} = (i, i+2)$ e iterando tale definizione un numero opportuno di volte si ottiene lo scambio

$$(i,j) = \tau_i\tau_{i+1}\dots\tau_{j-2}\tau_{j-1}\tau_{j-2}\tau_{j-3}\dots\tau_i\,.$$

Osservando che $(i,j) = (j,i)$, si ha l'asserto per *ogni* scambio con $j \neq i+1$. Si osservi a questo punto che ogni ciclo $(i_1, i_2, \dots, i_r)$ può essere ottenuto come prodotto di $(r-1)$ scambi nel modo seguente

$$(i_1, i_2, \dots, i_r) = (i_1, i_r)(i_1, i_{r-1})\dots(i_1, i_2)$$

e dunque anche come prodotto dei τ. Usando infine il fatto che ogni permutazione è ottenibile come prodotto di cicli, si ottiene che ogni permutazione σ è ottenibile come prodotto

$$\sigma = \tau_{i_1}\tau_{i_2}\dots\tau_{i_k}$$

dei generatori τ.

Si osservi che la procedura sopra descritta per ottenere ogni permutazione come prodotto dei generatori τ non garantisce affatto l'unicità di tale decomposizione (si trovi un esempio) e neppure il fatto che tale decomposizione sia di cicli a due a due disgiunti. Rimane dunque il problema di come determinare il prodotto di due permutazioni, *solo sapendo come si comporta il prodotto sui generatori.*

A questo scopo, procediamo come segue. Consideriamo l'insieme W di tutte le k-uple ordinate di generatori τ ("*parole di lunghezza k*"; si veda l'esempio 1.14.4)

$$w = \langle \tau_{i_1}, \tau_{i_2}, \dots, \tau_{i_k} \rangle\,,$$

dove gli indici $i_1, \dots i_k$ sono tutti compresi tra 1 e $n-1$, per qualsiasi naturale k. Introduciamo la seguente relazione: data una coppia di parole w e w', diciamo che "*w si riduce a w'*", e scriveremo

$$w \Rightarrow w'\,,$$

se w' si ottiene da w applicando una delle seguenti regole

(R_1) $\quad \langle \tau_i, \tau_i \rangle \Rightarrow 1$

(R_2) $\quad \langle \tau_i, \tau_j \rangle \Rightarrow \langle \tau_j, \tau_i \rangle\,,$ se $j < i-1$

(R_3) $\quad \langle \tau_i, \tau_{i-1}, \dots, \tau_j, \tau_i \rangle \Rightarrow \langle \tau_{i-1}, \tau_i, \tau_{i-1}, \dots, \tau_j \rangle\,,$ se $j < i$

alle sottoparole di w formate da lettere consecutive ("*segmenti*" di w). Diciamo che una parola w è in *forma normale* se *non è possibile ridurla ulteriormente* e osserviamo che *per ogni parola w esiste una catena di riduzioni*

$$w \Rightarrow w_1 \Rightarrow w_2 \Rightarrow \ldots \Rightarrow w_h$$

tali che w_h è in forma normale. Infatti, o w è già in forma normale, o appaiono due indici consecutivi cui si può applicare una delle regole R_1 o R_2, oppure tre o più indici consecutivi cui si può applicare la R_3. Se applichiamo R_1 la parola w_1 che otteniamo diminuisce di lunghezza, mentre se applichiamo una delle altre due regole, la parola che otteniamo *precede w nell'ordine lessicografico* (si veda 1.13.6). È chiaro che possiamo continuare ad applicare tale procedimento solo un numero finito di volte, perchè la lunghezza decresce e perchè le parole della stessa lunghezza che precedono una parola data nell'ordine lessicografico sono un numero finito e dunque che in un numero finito di passi troviamo una forma normale.

La questione cruciale è ora la seguente. È chiaro che il procedimento descritto per trovare una forma normale non è unico. Ad esempio posso cominciare ad esaminare le coppie o le terne di indici consecutivi cui applicare le regole a partire dall'inizio della parola in esame, oppure a partire dal fondo. In generale si otterranno due catene diverse di riduzioni e a questo punto noi sappiamo solo che entrambe devono condurre ad una forma normale. Il punto è ora di mostrare che due qualsiasi catene di riduzioni che partono dalla stessa parola, conducono sempre alla *stessa* forma normale o, in altre parole, che ogni parola può essere ridotta ad una *unica* forma normale. Questa proprietà è una proprietà assai forte sulle riduzioni, come mostra il seguente argomento. Sia

$$|\ \ |: \mathrm{W} \longrightarrow \mathrm{S}_n$$

la funzione che ad ogni parola w di generatori assegna il loro prodotto $|w|$ in S_n. Poichè, come abbiamo già osservato, in S_n le relazioni date sui generatori sono vere, si ha che se $w \Rightarrow w'$, allora $|w| = |w'|$. Questo è evidente per le riduzioni R_1 e R_2, mentre per la R_3 si ragiona per induzione sull'indice j nel modo seguente. L'enunciato vale se $j = i-1$, perchè è la terza delle relazioni nell'enunciato del teorema; supponiamo

dimostrata l'relazione sui generatori corrispondente alla regola R_3 e proviamo la corrispondente relazione per $j-1$:

$$\tau_i\tau_{i-1}\dots\tau_j\tau_{j-1}\tau_i = \tau_i\tau_{i-1}\dots\tau_j\tau_i\tau_{j-1} = \tau_{i-1}\tau_i\tau_{i-1}\dots\tau_j\tau_{j-1},$$

usando l'ipotesi di induzione e la seconda delle relazioni nell'enunciato. Dunque, se una parola avesse due forme normali, si potrebbe dimostrare esistono altre relazioni sui generatori che sono valide in S_n, ma che non sono tra quelle elencate e si potrebbe cominciare a dubitare del fatto che le relazioni date siano sufficienti a ricostruire il prodotto in S_n.

Per raggiungere il nostro scopo sarà sufficiente *contare* le forme normali e mostrare che esse sono proprio tante quanti gli elementi di S_n, cioè $n!$. Infatti, poichè ogni parola ha una forma normale, se una parola w avesse due forme normali, allora alla permutazione $|w|$ corrisponderebbero due forme normali, e dunque la cardinalità dell'insieme $\mathcal{F}_n$ delle forme normali non potrebbe essere quella di S_n. Inoltre, il fatto che dimostriamo che ogni parola ha un'unica forma normale solo usando le relazioni descritte nell'enunciato del teorema, garantisce che possiamo ricostruire il prodotto in S_n solo conoscendo tali relazioni: date due permutazioni, consideriamo le forme normali associate ad una loro decomposizione come prodotto dei generatori e consideriamo la forma normale della loro concatenazione; questa è la forma normale del prodotto, che evidentemente si ottiene solo usando le relazioni descritte nel teorema.

Cominciamo con il semplificare la notazione identificando una parola $w = \langle \tau_{i_1}, \tau_{i_2}, \dots, \tau_{i_k}\rangle$ semplicemente con la successione dei suoi indici $\langle i_1, i_2, \dots, i_k\rangle$, che altro non è che una funzione $i\colon [k] \longrightarrow [n-1]$. Così ad esempio per $n = 5$ la successione $\langle 2,3,1,4,2,1,3\rangle$ denota la parola $\langle \tau_2, \tau_3, \tau_1, \tau_4, \tau_2, \tau_1, \tau_3\rangle$ di lunghezza 7 (quale è una sua forma normale?). Per contare le forme normali, usiamo la tecnica standard di metterle in corrispondenza biunivoca con un insieme di cui sappiamo contare la cardinalità. Dato un qualsiasi naturale non nullo m, poniamo

$$\widetilde{[m]} = \{0, 1, 2, \dots, m\}$$

e sia F_n l'insieme

$$F_n = \{\rho\colon [n-1] \longrightarrow \widetilde{[n-1]} \mid \rho(i) \leq i, \text{ per ogni } i \in [n-1]\}\,.$$

Vogliamo mostrare che è possibile porre l'insieme F_n in corrispondenza biunivoca con l'insieme delle forme normali. Data $\rho \in \mathrm{F}_n$, definiamo una parola $w(\rho)$ nel modo seguente: dapprima definiamo per ogni $i \in [n-1]$ una parola $w(\rho(i))$ come

$$w(\rho(i)) = \begin{cases} \text{parola vuota} & \text{se } \rho(i) = 0 \\ \text{parola definita dalla successione} & \\ \langle i, i-1, \ldots, i-\rho(i)+1 \rangle & \text{se } \rho(i) \neq 0\,. \end{cases}$$

Definiamo quindi $w(\rho)$ come la parola che si ottiene concatenando le parole $w(\rho(1)), w(\rho(2)), \ldots, w(\rho(n-1))$ in questo ordine. Ad esempio, se $\rho\colon [4] \longrightarrow [4]$ è la funzione $\rho(1) = 0,\ \rho(2) = 2,\ \rho(3) = 3,\ \rho(4) = 1$, la parola $w(\rho)$ è la parola associata alla successione $\langle 2, 1, 3, 2, 1, 4 \rangle$. Si osservi che la funzione ρ soddisfa la clausola richiesta per cui $\rho(i) \leq i$. Per raggiungere il nostro scopo dobbiamo mostrare tre fatti:

1. $w(\rho)$ è in forma normale;
2. ogni parola in forma normale proviene da un'unica $\rho \in \mathrm{F}_n$ mediante il procedimento descritto (e dunque la definizione di $w(\rho)$ fornisce una corrispondenza biunivoca $w\colon \mathcal{F}_n \longrightarrow \mathrm{F}_n$ tra l'insieme delle forme normali e l'insieme F_n);
3. F_n ha cardinalità $n!$.

Cominciamo con il primo. Basta esaminare i segmenti di $w(\rho)$ che provengono da successioni della forma

$$\langle i - r - \rho(i-r), i, i-1, \ldots, i-\rho(i)+1, i+s \rangle\,,$$

dove $i-r$ è il più grande indice minore di i tale che $\rho(i-r) \neq 0$ e $i+s$ è il più piccolo indice maggiore di i tale che $\rho(i+s) \neq 0$. Le possibili riduzioni possono comparire solo nel segmento iniziale e in quello finale. Sia $d = i - (i - r - \rho(i-r) + 1) = r + \rho(i-r) - 1 \geq r - 1$. Dunque, se $r > 1$, allora $d > 0$, mentre se $r = 1$, allora $d = \rho(i-1)$, che in ogni caso è non nullo, perchè $i - r$ è il più grande intero minore di i per cui $\rho(i-r)$ è non nullo. Dunque il precedente di i è sempre minore di i e non si applicano le regole. L'argomento riguardante il segmento finale è del tutto simmetrico. Dunque le parole definite dalle successioni del tipo $w(\rho)$ sono in forma normale.

Per quanto riguarda il secondo fatto, sia

$$\langle i_1, i_2, \ldots, i_r, \ldots, i_k \rangle$$

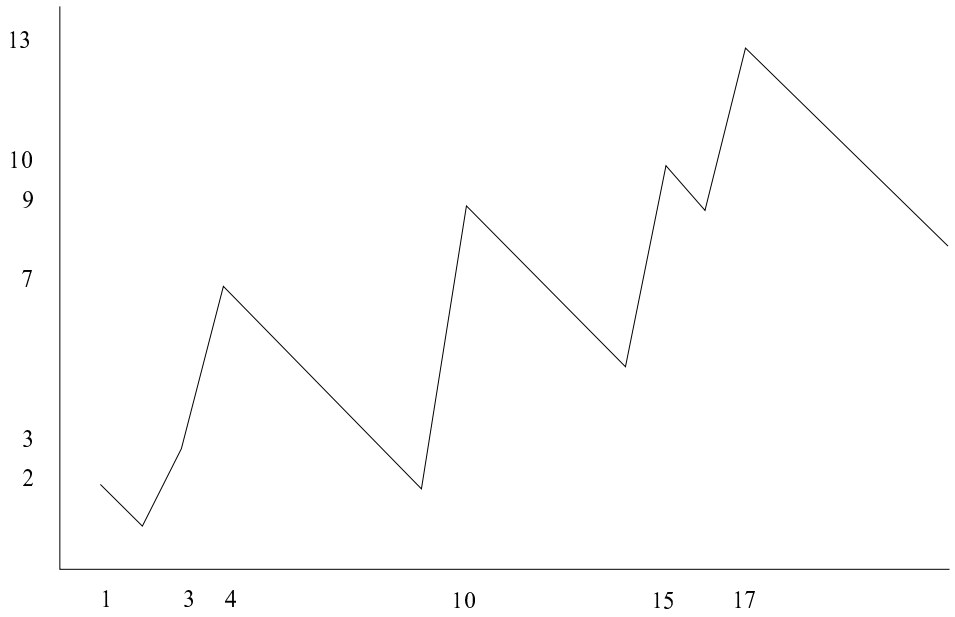
la successione degli indici di una parola w in forma normale e cerchiamo di capire come deve essere fatta o, meglio, come deve essere fatto il suo grafico. Esaminando da questo punto di vista le regole, e traducendole in regole di trasformazione del grafico di una spezzata che congiunge punti del piano a coefficienti interi, si può vedere che la successione degli indici $i\colon [k] \longrightarrow [n-1]$ di una parola in forma normale definisce una spezzata congiungente punti a del piano a coordinate intere se e solo se ha le seguenti proprietà:

1. la spezzata non ha tratti orizzontali;

2. i tratti discendenti sono paralleli alle bisettrice del primo e terzo quadrante;

3. la sottosuccessione dei valori assunti da i che maggiorano il valore che li precede è strettamente crescente.

Si osservi che in particolare la clausola 3 implica che i due segmenti uscenti da un punto di minimo relativo formano un angolo inferiore ad un angolo retto. Per meglio comprendere tale definizione, illustriamola con un esempio:



Certamente il grafico di una parola $w(\rho)$ ha queste proprietà. Viceversa, se il grafico di una parola i ha queste proprietà, allora la parola è in forma normale (basta esprimere le regole come regole di trasformazione di grafici di spezzate). In tal caso, se $1 \leq i_{h_1} < i_{h_2} < \ldots < i_{h_r} \leq (n-1)$ è la sottosuccessione dei valori assunti da i che maggiorano il precedente valore (convenendo che $h_1 = 1$), definiamo $\rho\colon [n-1] \longrightarrow \widetilde{[n-1]}$ come

$$\rho(j) = \begin{cases} 0 & \text{se } j \notin \{i_{h_1}, i_{h_2}, \ldots, i_{h_t}\} \\ i_{h_u} - i_{h_{u+1}-1} + 1 & \text{se } j = i_{h_u},\ u = 1, \ldots, t\,. \end{cases}$$

Con un po' di pazienza il lettore può verificare che per tale funzione ρ si ha $w(\rho) = i$ e che è l'unica possibile con tale proprietà.

Infine, il fatto che la cardinalità di F_n è $n!$ si prova con l'usuale tecnica di conteggio delle funzioni tra insiemi finiti, pur di tenere conto della clausola $\rho(i) \leq i$. ■

Prima di procedere è opportuno qualche commento sulla dimostrazione del teorema, in realtà alquanto laboriosa. La ragione di avere scelto questa dimostrazione (ne esistono di più semplici) è quella di cogliere l'opportunità di illustrare su un esempio concreto e rilevante alcune importanti idee che stanno alla base della Matematica Computazionale. Il problema è il seguente: esistono dimostrazioni di enunciati del tipo di quello del teorema precedente, che in principio possano essere automatizzate, cioè eseguite da una macchina? La risposta è affermativa e la dimostrazione precedente ne è un esempio. È chiaro che a partire dalle regole di riduzione (chiamate anche "*regole di riscrittura*"), il procedimento di trovare una forma normale può essere descritto da un algoritmo (cioè da un programma eseguibile da una macchina), pur di operare una scelta arbitraria sull'ordine con cui applicare le regole di riduzione alle parole. Primo problema: cosa mi assicura che per qualsiasi scelta l'algoritmo termini sempre? Ancora, cosa mi assicura che, ammesso che termini, il risultato non dipenda da queste scelte arbitrarie? Evidentemente queste sono proprietà dell'ordine indotto sull'insieme delle parole dalle regole di riduzione. Un altro problema cruciale è il seguente. Il lettore avrà notato che le regole di riduzione non sono semplicemente le relazioni dell'enunciato in cui si è sostituita l'uguaglianza con una freccia, ma comprendono anche altre conseguenze di quelle relazioni, in cui l'uguaglianza è sostituita con una freccia. La questione è dunque se, a partire da un certo numero finito di relazioni, è meccanizzabile il processo per trovare un insieme finito di regole di riduzione per cui i due problemi precedenti abbiano risposta affermativa. Ci limitiamo qui a queste prime considerazioni, peraltro basilari, rimandando ad un corso appropriato i lettori cui questa discussione abbia stimolato l'interesse per tali argomenti. Appare tuttavia doveroso almeno citare i nomi degli algoritmi maggiormente noti per questi problemi, che sono l'algoritmo di Knuth-Bendix, quello di Todd-Coxeter e, nel caso lineare, quello delle basi di Gröbner.

Un corollario immediato del teorema precedente è il seguente. Sia σ una permutazione qualsiasi e decomponiamola nel prodotto di simmetrie τ. Sappiamo che tale decomposizione non è unica, ma sappiamo anche che ogni due decomposizioni di σ si ottengono l'una dall'altra mediante applicazioni delle relazioni enunciate nel teorema, poichè hanno

la stessa forma normale. Ora, osservando la forma delle relazioni, si può notare facilmente che ogni loro applicazione lascia invariato il numero delle occorrenze dei generatori τ o lo altera di due. Dunque, la *parità* del numero delle occorrenze dei generatori in ogni decomposizione di σ *è sempre la stessa.* Dunque la definizione

$$\mathrm{sgn}(\sigma) = \begin{cases} 0 & \text{se } \sigma \text{ può essere decomposto in un numero pari di } \tau \\ 1 & \text{altrimenti} \end{cases}$$

è ben posta ed è immediato vedere che definisce un *omomorfismo* di gruppi

$$\mathrm{sgn}\colon \mathrm{S}_n \longrightarrow \mathbf{Z}_2,$$

il cui nucleo è il sottogruppo normale

$$\mathrm{A}_n = \ker(\mathrm{sgn})$$

formato dalle permutazioni "*pari*", quelle cioè che possono essere decomposte nel prodotto di un numero pari di generatori. Il gruppo A_n è detto gruppo "*alterno*" e ha ordine $|\mathrm{A}_n| = \frac{n!}{2}$. Conviene mettere in evidenza che componendo l'omomorfismo sgn con l'isomorfismo $\mathbf{Z}_2 \longrightarrow \{-1, 1\} = \mathbf{Z}^*$ si ottiene un omomorfismo

$$\epsilon\colon \mathrm{S}_n \longrightarrow \mathbf{Z}^*,$$

di uso frequente in molte parti della matematica.

Per finire, si osservi che se assegnamo un elemento $\phi(\tau_i) = g_i \in G$ ad ognuno degli $n-1$ generatori di S_n, con la condizione che *gli elementi g_i soddisfino in G le stesse relazioni cui soddisfano i generatori τ_i in S_n*, allora si può estendere tale definizione ad un unico omomorfismo di gruppi $\phi\colon \mathrm{S}_n \longrightarrow \mathrm{G}$. Tale osservazione permette spesso facilitare la definizione di omomorfismi di dominio S_n; ad esempio, la si può usare per provare quanto asserito nell'esempio 3 di 4.1.

4.5 Esercizi

1. Si provi che i cicli di lunghezza k di S_n sono precisamente le rappresentazioni fedeli del gruppo $\mathbf{Z}_k$ in S_n.

2. Si provi che due permutazioni sono disgiunte se e solo se $[\mathrm{Fix}(\sigma)]^c \subseteq \mathrm{Fix}(\tau)$ (si ricordi che A^c denota il complemento di un sottoinsieme $A \subseteq X$).

3. Dimostrare che il coniugato $\rho\gamma\rho^{-1}$ di un ciclo γ lunghezza k mediante una qualsiasi permutazione ρ è ancora un ciclo di lunghezza k.

4. Dimostrare che un ciclo è pari se e solo se il suo ordine è dispari. Dedurre un metodo per calcolare la parità di ogni permutazione.

5. Per ogni insieme ordinato X esistono due sottoinsiemi canonici del monoide $\mathrm{End}(X)$ delle endofunzioni di X: il sottogruppo $\mathrm{Aut}(X)$ degli automorfismi dell'insieme X e l'insieme $X!$ delle endofunzioni f di X con la proprietà $f(x) \leq x$, per ogni $x \in X$. Quando X è l'ordinale $[n]$, i due sottoinsiemi hanno la stessa cardinalità $n!$ e dunque sono isomorfi. Si utilizzi il teorema 4.4.2 per descrivere un esplicito isomorfismo (ricorsivo) $\Theta\colon [n]! \longrightarrow \mathrm{Aut}([\mathrm{n}]) = \mathrm{S}_n$, osservando dapprima che la definizione $\Psi(f)(i)) = f(i+1) - 1$ fornisce un isomorfismo $\Psi\colon [n]! \longrightarrow \mathrm{F}_n$.

6. Dimostrare che per ogni permutazione σ di S_n diversa dall'identità, $n > 2$, esiste uno scambio τ tale che $\sigma\tau \neq \tau\sigma$. Dedurre che per $n > 2$, il centro di S_n è costituito dalla sola identità.

7. Si dimostri che in S_4 ci sono esattamente 9 elementi di ordine 2, 8 di ordine 3 e 6 di ordine 4. Si deduca che S_4 ha 4 sottogruppi di ordine 3 e che i sottogruppi di ordine 6 sono tutti isomorfi a S_3.

8. Si provi che il gruppo alterno A_4 non ha sottogruppi di ordine 6.

4.6 I teoremi di Sylow

L'ultimo esercizio della precedente sezione mostra che il Teorema di Lagrange non può essere invertito: esiste un gruppo finito e un divisore dell'ordine del gruppo per cui non esistono sottogruppi aventi per ordine quel divisore. I teoremi di Sylow si preoccupano di trovare condizioni su un divisore k dell'ordine n di un gruppo finito G per le quali è

garantita l'esistenza di un sottogruppo di quell'ordine. La risposta è alquanto sorprendente ed è legata in modo ancora un po' misterioso al teorema fondamentale della aritmetica: basta che k sia della forma p^α, essendo p uno dei primi in cui n si decompone n in modo unico come prodotto di potenze di primi e α un naturale non superiore alla potenza con cui p compare nella decomposizione di n. Esistono poi altri due teoremi che danno ulteriori informazioni sul loro numero e sulle loro relazioni, quando α è la massima potenza di p per cui p^α divide n.

Teorema 4.6.1 *(Primo Teorema di Sylow) Se p è un primo e se p^α è un divisore dell'ordine n di un gruppo finito* G*, allora* G *ha un sottogruppo di ordine p^α.*

DIMOSTRAZIONE. Posto $n = p^\alpha m$, sia $\mathbf{P}_{p^\alpha}G$ l'insieme dei sottoinsiemi di G aventi p^α elementi e consideriamo l'azione di G su $\mathbf{P}_{p^\alpha}G$ indotta dal prodotto (si veda l'esempio 4.1.5). Sia r l'unico naturale tale che p^r divide m, ma p^{r+1} non divide m e mostriamo che esiste un'orbita di tale azione la cui cardinalità non è multipla di p^{r+1}. Infatti, osservando dapprima che p^r è un divisore della cardinalità $\binom{p^\alpha m}{p^\alpha}$ di $\mathbf{P}_{p^\alpha}G$, ma che p^{r+1} non lo è[1], se p^{r+1} dividesse la cardinalità di ogni orbita, dato che le orbite sono una partizione di $\mathbf{P}_{p^\alpha}G$, dividerebbe anche la sua cardinalità, assurdo. Sia dunque S *un sottoinsieme di G di cardinalità p^α la cui orbita ha una cardinalità k che non è un multiplo di p^{r+1}* e sia

$$\mathrm{H} = \mathrm{st}(S)$$

il sottogruppo di G dato dal suo stabilizzatore. Mostreremo che la cardinalità $|H|$ di H è p^α. Sappiamo che la cardinalità di un'orbita è l'indice dello stabilizzatore di un qualsiasi suo elemento (si veda 4.3), dunque

$$k = \frac{p^\alpha m}{|H|},$$

[1] Infatti

$$\binom{p^\alpha m}{p^\alpha} = \frac{p^\alpha m(p^\alpha m - 1)\dots(p^\alpha m - (p^\alpha - 1))}{p^\alpha(p^\alpha - 1)!} = m\prod_{i=1}^{p^\alpha - 1}\frac{(p^\alpha m - i)}{p^\alpha - i}$$

e se $p^t \leq p^\alpha$, allora p^t divide $p^\alpha - i$ se e solo se divide $p^\alpha m - (p^\alpha - i)$; dunque in $\binom{p^\alpha m}{p^\alpha}$ tutte le potenze p^t con $t \leq \alpha$ scompaiono, tranne quelle che dividono m.

cioè $k|H| = p^\alpha m$. Poichè p^{r+1} non divide k e $p^{\alpha+r}$ divide $p^\alpha m = k|H|$, si ha che p^α divide $|H|$ e quindi anche $p^\alpha \leq |H|$. Infatti, poichè $p^{\alpha+r}$ divide $k|H|$, *per la unicità della decomposizione in prodotto di potenze di primi di* $k|H|$, si ha che una potenza p^t, con $\alpha + r \leq t$, divide $k|H|$. Sia $t = u+v$, tali che p^udivide k e p^v divide $|H|$; poichè p^{r+1} non divide k, deve essere $u \leq r$, dunque $\alpha + u \leq \alpha + r \leq u + v$, dunque $\alpha \leq v$; perciò p^α divide p^v, che divide $|H|$. D'altra parte, fissato $s \in S$, per ogni elemento $g \in H$ il prodotto sg è ancora un elemento di S, *per come è definito* H; dunque tale moltiplicazione definisce una funzione $H \longrightarrow S$ che è *iniettiva* (basta usare l'inverso s^{-1}) e quindi $|H| \leq |S| = p^\alpha$; perciò

$$|H| = p^\alpha \,. \blacksquare$$

Se p^α è la *massima* potenza di un primo p che divida l'ordine di un gruppo finito G, un sottogruppo di ordine p^α di G di cui abbiamo appena mostrato l'esistenza si dice "*p-gruppo (di Sylow).* È chiaro che il coniugato di ogni p-gruppo di Sylow di G è ancora tale. Il secondo teorema di Sylow inverte tale affermazione. Per la dimostrazione dobbiamo premettere alcune considerazioni:

1. Dati due gruppi G e H, chiameremo "*(*G-H*)-insieme*" una azione $(\mathrm{G}\times\mathrm{H}^{op})\times X \longrightarrow X$ del gruppo $\mathrm{G}\times\mathrm{H}^{op}$ su un insieme X. L'esempio più semplice si ottiene prendendo $G = H = X$ e definendo l'azione mediante $\langle a, b\rangle x = axb$. Il lettore è invitato a verificare gli assiomi. È chiaro che se A e B sono sottogruppi di G, tale azione si restringe ad un (A-B)-insieme. Come per ogni azione, le orbite costituiscono una partizione, che nell'esempio ora descritto di G come (A-B)-insieme si chiamano "*laterali doppi*" e vengono denotati con AxB.

2. Se A e B sono sottoinsiemi di un gruppo G, definiamo il prodotto AB come il sottoinsieme di G

$$AB = \{ab \,|\, a \in A \ e \ b \in B\} \,.$$

Anche se A e B sono sottogruppi, non è detto che AB lo sia: lo è se e solo se $AB = BA$ e in tal caso è il più piccolo sottogruppo che li contiene entrambi, che denoteremo con $\mathrm{A} \vee \mathrm{B}$. La semplice dimostrazione è lasciata per esercizio al lettore. Inoltre, se A e B

sono sottogruppi di un gruppo finito G, si ha la seguente formula sulla cardinalità del prodotto AB:

$$|AB| = \frac{|A||B|}{|A \cap B|}\,.$$

Basta infatti mostrare che ogni elemento ab di AB ha $|A \cap B|$ scritture distinte. Se $t \in A \cap B$, allora $ab = (at)(t^{-1}b)$, con $at \in A$ e $t^{-1}b \in B$; se $ab = uv$, allora $u^{-1}a = vb^{-1} = t \in A \cap B$.

Teorema 4.6.2 *(Secondo Teorema di Sylow) Due p-gruppi di Sylow di un gruppo finito* G*, relativi allo stesso primo p, sono coniugati.*

Dimostrazione. Osserviamo che se A e B sono sottogruppi di G, allora la moltiplicazione con x^{-1} definisce una corrispondenza biunivoca $AxB \longrightarrow AxBx^{-1}$; poichè xBx^{-1} è un sottogruppo di G isomorfo a B, per la precedente osservazione 2 si ha:

$$|AxB| = |AxBx^{-1}| = \frac{|A||xBx^{-1}|}{|A \cap xBx^{-1}|} = \frac{|A||B|}{|A \cap xBx^{-1}|}\,.$$

Siano ora A e B due p-gruppi di Sylow di G. Procediamo per assurdo. Se per ogni $x \in G$ si avesse $A \neq xBx^{-1}$, allora per ogni x si avrebbe $|A \cap xBx^{-1}| = p^m$ con $m < n$, dove n è il massimo naturale per cui p^n divide l'ordine di G. Dunque

$$|AxB| = \frac{p^{2n}}{p^m} = p^{2n-m}$$

e $2n-m \geq n+1$, poichè $2n = n+n \geq m+n+1$. Di qui la contraddizione: se per ogni $x \in G$, p^{n+1} dividesse $|AxB|$, allora p^{n+1} dividerebbe $|G|$, perchè i laterali doppi AxB sono una partizione di G. Dunque esiste $x \in G$ tale che $A = xBx^{-1}$. ■

Per quanto riguarda il terzo Teorema di Sylow, ricordiamo dapprima che per l'azione di coniugio sui sottogruppi $G \times Sub(\mathrm{G}) \longrightarrow Sub(\mathrm{G})$ lo stabilizzatore st(H) di un sottogruppo H di G

$$\mathrm{st(H)} = \{g \in G \,|\, gHg^{-1} = H\}$$

è un sottogruppo di G che contiene H, detto "normalizzante" di H e denotato con N(H. Sappiamo che l'indice di N(H in G è il numero dei coniugati distinti, che è il numero degli elementi dell'orbita di H nella azione di coniugio sui sottogruppi.

Teorema 4.6.3 *(Terzo Teorema di Sylow) Sia P un p-sottogruppo di Sylow di un gruppo finito* G. *Il numero* $\frac{|G|}{|N(P|}$ *dei coniugati distinti di* P, *dunque il numero dei p-sottogruppi di Sylow di* G, *è della forma*

$$\frac{|G|}{|N(P|} = 1 + kp\,.$$

DIMOSTRAZIONE. Si ha

$$|PxP| = \frac{|P|^2}{|P \cap xPx^{-1}}\,.$$

Dunque, se $x \notin N(\mathrm{P})$, cioè se $P \cap xPx^{-1} \neq P$, allora p^{n+1} non divide $|PxP|$, per n tale che $|P| = p^n$, poichè $|PxPx^{-1} = p^m$ con $m < n$ e quindi $|PxP| = p^{2n-m} \geq p^{n+1}$. Se invece $x \in N(\mathrm{P})$, cioè $P = xPx^{-1}$, allora $Px = xP$ e $PxP = P^2x = Px$ e dunque $|PxP| = p^n$. Riassumendo, $x \in N(\mathrm{P})$ se e solo se $|PxP| = p^n$. Consideriamo ora l'equazione delle classi

$$|G| = \sum_{x \in N(\mathrm{P})} |PxP| + \sum_{x \notin N(\mathrm{P})} |PxP|$$

dove la somma è estesa ad una scelta arbitraria di elementi $x \in G$, uno ed uno solo per ogni laterale doppio. Se $x \in N(\mathrm{P})$, allora $PxP = Px$ e dunque la prima somma è $|N(\mathrm{P})|$, perchè la somma è estesa ad una scelta di x, uno per ogni laterale di P in $N(\mathrm{P})$. Ogni termine della seconda somma è divisibile per p^{n+1}, dunque p^{n+1} la divide e perciò esiste un numero razionale u tale che

$$\sum_{x \notin N(\mathrm{P})} |PxP| = up^{n+1}\,.$$

Perciò

$$|G| = |N(\mathrm{P})| + up^{n+1}$$

e dunque

$$\frac{|G|}{|N(\mathrm{P}|} = 1 + u\frac{p^{n+1}}{|N(\mathrm{P})|}\,.$$

Ma $|N(\mathrm{P})|$ divide $|G|$, poichè $N(\mathrm{P})$ è un sottogruppo di G. Dunque $u\frac{p^{n+1}}{|N(\mathrm{P})|}$ deve essere *intero*. Inoltre p^{n+1} non divide $|N(\mathrm{P})|$, perchè se lo dividesse, allora dividerebbe $|G|$, dato che $|N(\mathrm{P})|$ divide $|G|$. Dunque $u\frac{p^{n+1}}{|N(\mathrm{P})|}$ deve essere divisibile per p, cioè della forma kp, con k intero. ■