

# Algebra

Logica rudimentale



# Logica rudimentale

GIOVANNI CUTOLO

<b>1. Premesse</b>	<b>1</b>	<b>5. Quantificatori</b>	<b>13</b>
<b>2. Connettivi proposizionali</b>	<b>2</b>	Variabili libere e vincolate	14
Negazione	2	Quantificatori ristretti	16
Congiunzione	2	<b>6. Qualche regola d'uso</b>	<b>17</b>
Tavole di verità	2	Quantificatori multipli	17
Disgiunzione (inclusiva)	3	Negazione di quantificatori	18
Equivalenza, o Doppia implicazione	4	Un errore da evitare	19
Implicazione	4	<b>7. Insiemi</b>	<b>19</b>
<b>3. Tautologie</b>	<b>6</b>	Formule insiemistiche	21
Alcune tautologie elementari	7	Ancora De Morgan	22
Distributività	8	Differenza simmetrica	22
Leggi di De Morgan	9	Parti di un fissato insieme	23
Tautologie sulla implicazione	9		
<b>4. Altri connettivi binari</b>	<b>11</b>		

## 1. PREMESSE

Lo scopo di queste note è quello di dare alcune indicazioni sull'uso corretto dei simboli logici utili, o piuttosto necessari, nello studio della matematica. Una vera e propria (e più rigorosa) introduzione alla logica va al di là degli obiettivi del corso di Algebra ed è rimandata a corsi degli anni successivi e della laurea magistrale.

Molto informalmente, ci limitiamo qui a dire che ogni teoria matematica è espressa in quello che si chiama un *linguaggio*, che è costituito (1) da un alfabeto di simboli (semplicemente, dei caratteri tipografici), che possono essere messi insieme per costituire parole (stringhe di caratteri) e (2) da regole *sintattiche* che permettano, tra l'altro, di distinguere tra stringhe “correttamente composte”, che si chiamano *formule* e stringhe che non sono correttamente composte. L'aggettivo ‘sintattiche’ indica che le regole riguardano solo le modalità di manipolazione formale dei simboli e non fanno alcun riferimento a ciò che questi simboli intendono rappresentare.

Spero che un esempio possa aiutare a chiarire la nozione di formula. Il linguaggio di una teoria che voglia descrivere l'usuale aritmetica dei numeri interi potrebbe contenere dei simboli per indicare variabili (ad esempio, “ $x$ ”, “ $y$ ”, “ $z$ ”, …), alcuni simboli che indichino costanti (come “0” e “1”), operazioni (ad esempio, “+”, “.”, “−”), relazioni (come “ $<$ ”, ma anche, su un piano diverso, “ $=$ ”). Le regole sintattiche saranno probabilmente scritte in modo che stringhe come “ $x + 1 = y$ ” oppure “ $x \cdot y < 1 + z$ ” siano formule (*ben formate*, come anche si usa dire), mentre “ $x <$ ”, “ $x == y$ ” oppure “ $x + y$ ” non lo siano.<sup>1</sup>

È utile ribadire che la sintassi prescinde completamente dall'interpretazione dei simboli utilizzati nel linguaggio, e va tenuta ben separata da questa interpretazione. Ad esempio, non è affatto detto che i simboli che appaiono nel linguaggio dell'aritmetica appena richiamato debbano davvero essere interpretati come i nostri abituali numeri interi, che “+” debba indicare la consueta addizione, che le costanti “0” e “1” rappresentino necessariamente i familiari numeri zero ed uno, e così via. La parte della logica che si occupa di queste “interpretazioni” si chiama semantica e, in una trattazione rigorosa, va tenuta sempre ben distinta dalla sintassi.

---

<sup>1</sup>Osserviamo di passaggio che le regole sintattiche di un linguaggio hanno anche lo scopo di distinguere, tra le possibili stringhe, i cosiddetti *termini*, che così come i simboli di variabile e di costante intendono rappresentare gli oggetti “di cui parla” la teoria; nel nostro esempio i numeri interi. Delle tre stringhe che non sono formule appena mostrate, le prime due non sono termini, l'ultima (“ $x + y$ ”) invece lo è.

Una delle nozioni semantiche fondamentali, che siamo abituati a dare per scontata, è quella di verità o falsità di una affermazione. Ad esempio, facendo ancora riferimento all'aritmetica, con le consuete interpretazioni dei simboli che qui appaiono, siamo abituati a considerare vera la formula “ $0 < 1$ ” e falsa la formula “ $0 + 0 = 1$ ”. Abbiamo qualche perplessità, invece, a proposito di “ $x > 1 + 1$ ” (dove  $x$  è una variabile); diremmo che il valore di verità (cioè se essa è vera o falsa) di questa formula *dipende* (qualsiasi cosa ciò significhi) da  $x$ ; quindi questa formula non ha un valore di verità nel senso più intuitivo.

Esiste una classe di formule alle quali, in modo piuttosto ragionevole, è sempre possibile in linea di principio, attribuire un valore di verità. Queste sono le cosiddette *formule chiuse*, che vengono anche chiamate *proposizioni*,<sup>2</sup> o sentenze (con una discutibile traduzione dell'inglese *sentences*).

Daremo solo [più avanti](#) una definizione di formula chiusa (e sarà comunque una definizione piuttosto approssimativa; la definizione precisa di formula chiusa richiede tecnicismi di cui in queste note è bene fare a meno). Per ora ci accontentiamo di sapere che sono chiuse tutte le formule ben formate in cui non appaiano variabili (ma, attenzione!, come vedremo esistono formule chiuse contenenti variabili). Delle tre formule esibite poco sopra, sono chiuse le prime due (“ $0 < 1$ ” e “ $0 + 0 = 1$ ”, che non contengono variabili) ma non la terza (“ $x > 1 + 1$ ”).

È interessante sapere che la nozione di formula chiusa è sintattica, non semantica; ad essere rigorosi avremmo dovuto introdurre questa nozione prima di quelle semantiche di interpretazione e di verità.

## 2. CONNETTIVI PROPOSIZIONALI

Ogni linguaggio contiene dei simboli, i cosiddetti simboli logici, che permettono di costruire formule a partire da formule più semplici. Tra i simboli logici appaiono di regola (alcuni dei) *connettivi proposizionali*, che presenteremo in questa sezione. Il nostro punto di vista sarà essenzialmente semantico: descriveremo i connettivi proposizionali guardando a come essi influenzano i valori di verità delle proposizioni in cui appaiono. È appena il caso di ripetere che ci stiamo discostando da quanto sarebbe richiesto da una trattazione rigorosa.

**Negazione.** Il connettivo più semplice da descrivere è quello di negazione:  $\neg$ , che si indica anche con NOT o talvolta con  $\sim$  e che possiamo semplicemente leggere come “non”. Se  $p$  è una formula allora  $\neg p$  (oppure  $\neg(p)$ ; anche le parentesi, usate come di consueto per suggerire come vadano raggruppati i simboli, sono spesso comprese nell'alfabeto di un linguaggio) è anch'essa una formula. Se  $p$  è una proposizione vera, allora  $\neg p$  sarà una proposizione falsa; se  $p$  è falsa, allora  $\neg p$  è vera. Detto in modo più sintetico, il connettivo di negazione ha come argomento una sola formula (questo fatto si esprime dicendo che la negazione è un connettivo *unario*) ed inverte il valore di verità del suo argomento.<sup>3</sup>

**Congiunzione.** Il connettivo di congiunzione si indica con  $\wedge$  (oppure con AND) ed è, come tutti quelli che definiamo di seguito, un connettivo *binario*, vale a dire: richiede due formule come argomenti. Se  $p$  e  $q$  sono proposizioni,  $p \wedge q$  è una proposizione che è vera quando sono vere sia  $p$  che  $q$ , falsa altrimenti. Dunque,  $\wedge$  corrisponde alla congiunzione “e” del linguaggio ordinario, e possiamo leggere  $p \wedge q$  come “ $p$  e  $q$ ”.

**Tavole di verità.** La semantica dei connettivi appena introdotti è sintetizzata, in modo molto efficace, da *tavole* (o *tabelle*) di verità:

negazione		congiunzione		
$p$	$\neg p$	$p$	$q$	$p \wedge q$
$V$	$F$	$V$	$V$	$V$
$F$	$V$	$V$	$F$	$F$
		$F$	$V$	$F$
		$F$	$F$	$F$

Vediamo di cosa si tratta. Ciascuna delle due tavole è divisa in due parti. La parte destra consiste di una colonna, la cui intestazione è una formula ( $\neg p$  per la prima tavola,  $p \wedge q$  per la seconda); in questa

<sup>2</sup>avvertenza: alcuni autori usano il termine ‘proposizione’ come sinonimo di ‘formula’, non necessariamente chiusa.

<sup>3</sup>ad essere più precisi, dovremmo aggiungere la condizione che il valore di verità dell’argomento sia definito.

colonna andrà letto il valore di verità della formula. Nelle formule che stiamo esaminando appaiono delle variabili ( $p$  e  $q$ ), che rappresentano proposizioni. Nella parte sinistra di ciascuna delle tavole abbiamo una colonna per ogni variabile che appare nella formula considerata (quindi solo una colonna, intestata da  $p$ , per la prima tavola; due colonne, intestate da  $p$  e  $q$ , per la seconda). Guardiamo alla prima tavola, quella della negazione. I valori di verità possibili per la variabile  $p$  sono ovviamente due: vero ( $V$ ) e falso ( $F$ ). In corrispondenza di questi due possibili valori abbiamo due righe: la prima ci dice (guardando la colonna destra) che la formula  $\neg p$  è falsa quando  $p$  è vera, la seconda che  $\neg p$  è vera quando  $p$  è falsa, in accordo con quanto avevamo detto definendo il connettivo  $\neg$ . La seconda tavola ha invece quattro righe, perché la formula  $p \wedge q$  ha due variabili e le combinazioni possibili per i valori di verità di due variabili sono quattro. Nell'ordine in cui appaiono nella tavola, le possibilità sono: (1)  $p$  e  $q$  sono entrambe vere; (2)  $p$  è vera e  $q$  è falsa; (3)  $p$  è falsa e  $q$  è vera; (4)  $p$  e  $q$  sono entrambe false. Anche qui ciascuna riga riporta, nella colonna di destra, il valore di verità della formula (in questo caso  $p \wedge q$ ) in funzione dei valori di verità di  $p$  e  $q$  che appaiono, nella stessa riga, a sinistra. Come si vede, la tavola fornisce, anche in questo caso, esattamente le stesse informazioni che avevamo dato come definizione (semantica) di  $\wedge$ .

Completiamo questa introduzione alle tavole di verità con un minimo di terminologia e qualche indicazione ulteriore. Le variabili (come  $p$  e  $q$ , nei nostri esempi) che rappresentano proposizioni vengono chiamate *variabili proposizionali*; le formule costituite da variabili proposizionali, connettivi proposizionali (i due già definiti e quelli che stiamo per definire) e parentesi si chiamano *forme proposizionali*. Il *calcolo proposizionale* è la parte della logica che studia le forme proposizionali.

Come si può facilmente immaginare, esistono tavole di verità più complesse delle due che abbiamo esibito. Incontreremo tavole di verità che descrivono contemporaneamente più forme proposizionali, e quindi hanno più di una colonna nella parte destra, non solo una come nei nostri esempi. Incontreremo anche tavole che descrivono forme proposizionali con più di due variabili. È utile verificare (esercizio!) e ricordare poi che la tavola di verità di una forma proposizionale con un numero  $k$  di variabili richiede  $2^k$  righe, perché  $2^k$  è il numero di possibili combinazioni di valori di verità per  $k$  variabili.

### Esercizi.

- A.1. Scrivere le tavole di verità di ciascuna delle forme proposizionali: “ $p \wedge p$ ”, “ $(\neg p) \wedge q$ ” e “ $(\neg p) \wedge (\neg q)$ ”.
- A.2. Scrivere le tavole di verità delle forme proposizionali “ $p \wedge (q \wedge r)$ ” e “ $p \wedge (q \wedge (\neg r))$ ”.

Riprendiamo ora nostra lista di connettivi proposizionali binari. Come nel caso di  $\wedge$ , anche per i successivi vale questa regola sintattica: se  $p$  e  $q$  sono formule e  $*$  un connettivo binario allora  $p * q$  è una formula. Inoltre, se  $p$  e  $q$  sono proposizioni allora  $p * q$  è una proposizione ed il suo valore di verità dipende solo dal connettivo  $*$  e dai valori di verità di  $p$  e  $q$ .

**Disgiunzione (inclusiva).** Il connettivo di *disgiunzione* si indica con  $\vee$  o con OR. Da un punto di vista formale la sua descrizione è altrettanto semplice che quella della congiunzione, ma nell'uso la disgiunzione presenta qualche difficoltà in più. La ragione è che questo connettivo corrisponde ad uno dei significati che la particella<sup>4</sup> “o” ha in italiano (ovvero, che “or” ha in inglese). Il problema è, appunto, che nel linguaggio corrente “o” può assumere più significati, che hanno valore logico molto diverso.<sup>5</sup> Il significato che viene attribuito al connettivo proposizionale  $\vee$  in matematica è quello, come si dice, *inclusivo*, per il quale, se  $p$  e  $q$  sono proposizioni, la forma  $p \vee q$  è vera a condizione che *almeno una* tra  $p$  e  $q$  sia vera, ed è falsa nell'altro caso, cioè quando sia  $p$  che  $q$  siano false. La tavola di verità che definisce questo connettivo è dunque:

---

<sup>4</sup>secondo la grammatica della lingua italiana questa ‘particella’ è una congiunzione, ma evitiamo di chiamarla così per non fare confusione con il connettivo proposizionale  $\wedge$ , che abbiamo chiamato congiunzione.

<sup>5</sup>la necessità di rimuovere questo genere di ambiguità che sono proprie del linguaggio naturale è uno dei motivi per i quali, nel discorso scientifico, si deve utilizzare un linguaggio almeno parzialmente formalizzato.

disgiunzione (inclusiva)

$p$	$q$	$p \vee q$
$V$	$V$	$V$
$V$	$F$	$V$
$F$	$V$	$V$
$F$	$F$	$F$

Mettiamo in evidenza che  $p \vee q$  risulta vera anche quando  $p$  e  $q$  sono entrambe vere. Invece, se il menu turistico della trattoria prevede una portata di carne *o* una di pesce, possiamo esser certi che l'oste non intende questo “*o*” come una disgiunzione inclusiva, ma come quella che si chiama una disgiunzione *esclusiva*: il cliente ha diritto ad avere un piatto di carne oppure uno di pesce, ma non entrambi. Sulla disgiunzione esclusiva (indicata con  $\dot{\vee}$ , oppure con XOR) torneremo [più avanti](#); per ora diciamo che i valori di verità di  $p \vee q$  e  $p \dot{\vee} q$  differiscono solo nel caso in cui  $p$  e  $q$  siano entrambe vere; in questo caso  $p \vee q$  è vera,  $p \dot{\vee} q$  è falsa. Come detto, in italiano (e in inglese, ed in altre lingue) purtroppo sia la disgiunzione inclusiva che quella esclusiva sono rese dalla stessa particella (“*o*”, “or” etc.), ma qualcuno probabilmente ricorda che in latino questi due connettivi sono ben distinti anche nel linguaggio ordinario: *vel* esprime la disgiunzione inclusiva, *aut* quella esclusiva.<sup>6</sup>

È importante insistere (da parte di chi scrive) e ancora di più ricordare (da parte di chi legge) che in matematica, e generalmente nel linguaggio scientifico, *per disgiunzione si intende sempre la disgiunzione inclusiva*. Confusione su questo punto porta invariabilmente a pericolose incomprensioni e macroscopici errori.

**Equivalenza, o Doppia implicazione.** Questo connettivo, chiamato anche *bicondizionale* e indicato con  $\Leftrightarrow$  oppure con  $\leftrightarrow$ , si potrebbe definire in termini dell'implicazione, che verrà descritta tra poco, ma è molto semplice da descrivere in modo diretto. Se  $p$  e  $q$  sono proposizioni, “ $p \Leftrightarrow q$ ” (che viene letta come “ $p$  se e solo se  $q$ ”, oppure “ $p$  equivale a  $q$ ”) è vera se  $p$  e  $q$  hanno lo stesso valore di verità, falsa altrimenti. La relativa tavola di verità è dunque:

equivalenza		
$p$	$q$	$p \Leftrightarrow q$
$V$	$V$	$V$
$V$	$F$	$F$
$F$	$V$	$F$
$F$	$F$	$V$

**Implicazione.** Questo connettivo è chiamato anche *condizionale* e si indica con  $\Rightarrow$  o con  $\rightarrow$ . Ancora più che la disgiunzione, esso presenta delle difficoltà dovute al fatto che l'uso che se fa in logica non combacia con l'uso suggerito dal linguaggio ordinario. Se  $p$  e  $q$  sono formule, la formula “ $p \Rightarrow q$ ”, o meglio la sua espressione verbale: “se  $p$  allora  $q$ ” nel linguaggio quotidiano presuppone che ci sia un qualche nesso tra  $p$  e  $q$  che faccia apparire  $q$  come conseguenza di  $p$ . Ad esempio, “se piove allora non esco di casa” appare una frase del tutto naturale: si intende che il fatto che io non esca di casa dipende proprio dalla pioggia. Invece, la frase “se piove allora il Vesuvio è alto più di mille metri sul livello del mare” sembra priva di senso, per chi non ha studiato logica. Chi invece lo ha fatto ed è disposto a trasportare al linguaggio quotidiano i modi di espressione del linguaggio formale potrà convenire che la frase, per quanto di genere poco usuale, presa alla lettera non solo un senso lo ha ma è addirittura vera, come stiamo per vedere.

La tavola di verità che definisce l'implicazione è:

<sup>6</sup>a titolo di (forse comunque istruttiva) curiosità menzionamo il fatto che “*o*” può assumere in italiano almeno un altro significato. Se le richieste del piccolo Pasqualino superano un certo livello di soglia, la mamma potrà cercare di arginarle dicendogli che può avere “*o* il gelato *o* la pizzetta”. Probabilmente in questo caso la mamma intende solo dire che non può avere entrambe le cose, ma che è più che soddisfatta se Pasqualino rinuncia ad entrambe. In questo caso, dunque, “*o*” esprime il connettivo **NAND** (negazione della congiunzione) a cui faremo cenno in seguito.

implicazione		
$p$	$q$	$p \Rightarrow q$
$V$	$V$	$V$
$V$	$F$	$F$
$F$	$V$	$V$
$F$	$F$	$V$

Dunque, la formula  $p \Rightarrow q$  (dove  $p$  e  $q$  sono proposizioni) è vera sempre tranne che nel caso in cui  $p$  (che si chiama *antecedente* dell'implicazione<sup>7</sup>) è vero e  $q$  (che si chiama *conseguente* dell'implicazione) è falso. Mentre più o meno tutti sono subito d'accordo con questa definizione nel caso in cui l'antecedente sia vero, quasi nessuno studente accetta di buon grado la seconda parte della definizione (data dalle ultime due righe della tavola di verità), cioè il fatto che la formula si consideri vera quando l'antecedente è falso (per giunta, indipendentemente dal valore di verità del conseguente!). Lo scopo dei prossimi paragrafi è quello di mostrare che questa definizione non solo non è frutto di una scelta capricciosa di qualcuno, ma è l'unica coerente con l'uso 'intuitivo' che dell'implicazione siamo sempre stati abituati a fare.

L'obiezione che spesso, in aula, gli studenti sollevano è che 'quando si sa già che  $p$  è falso la frase "se  $p$  allora  $q$ " (o la formula " $p \Rightarrow q$ ") non ha senso', e quindi non dovrebbe essere definito il suo valore di verità. E, inoltre, molti ritengono che se anche si deve proprio attribuire a questa frase un valore di verità sia innaturale la scelta di porla vera.

Non è così: una delle caratteristiche specifiche dei linguaggi formalizzati è, per così dire, l'indifferenza rispetto ai contenuti. La correttezza sintattica di una formula (come " $p \Rightarrow q$ "), ed il fatto che essa venga interpretata nella semantica (in parole semplici, il fatto che la formula 'abbia senso') deve poter essere stabilita una volta per tutte e non può dipendere da informazioni accessorie di cui possiamo disporre o meno, come il contenuto fattuale di  $p$  (e che sono magari soggette a cambiare col tempo e le circostanze<sup>8</sup>). Questo fatto è addirittura utile, altrimenti, a rigore, non potremmo neanche discutere di formule come " $p \Rightarrow q$ " se avessimo il dubbio che si possa dimostrare la falsità di  $p$ ; ad essere pignoli non sarebbero possibili in matematica le dimostrazioni per assurdo o per contrapposizione.

Detto ciò, perché le implicazioni con antecedente falso devono proprio essere vere? Spero che questo esempio chiarisca la ragione della definizione. Consideriamo la frase "per ogni numero intero  $x$  compreso tra 1 e 3 si ha che se  $x > 2$  allora  $x > 1$ ". Tutti concordiamo su fatto che questa frase è vera. Analizziamola; essa significa che tutte le implicazioni  $x > 2 \Rightarrow x > 1$  ottenute sostituendo ad  $x$  uno dei numeri 1, 2, 3 sono vere. Dunque sono vere la proposizione  $\Phi_1$  : " $1 > 2 \Rightarrow 1 > 1$ ", la  $\Phi_2$  : " $2 > 2 \Rightarrow 2 > 1$ " e la  $\Phi_3$  : " $3 > 2 \Rightarrow 3 > 1$ ". Vediamo che la  $\Phi_3$  è del tipo 'non contestato': una implicazione in cui sia antecedente che conseguente sono veri, quindi vera, come indica il primo rigo della nostra tavola di verità. Le altre due proposizioni hanno invece l'antecedente falso. La  $\Phi_1$  è del tipo descritto dal quarto rigo della tavola di verità ('falso implica falso'), dal momento che sia l'antecedente ( $1 > 2$ ) che il conseguente ( $1 > 1$ ) sono falsi; la  $\Phi_2$  è del tipo descritto dal terzo rigo ('falso implica vero'). Se siamo d'accordo che sia ragionevole sostenere che la frase da cui siamo partiti ("per ogni numero intero  $x$  compreso tra 1 e 3 si ha che se  $x > 2$  allora  $x > 1$ ") sia vera, allora dobbiamo essere d'accordo anche sul fatto che siano (ragionevolmente) vere  $\Phi_1$  e  $\Phi_2$ , quindi che siano ragionevoli i valori di verità che appaiono nella nostra tabella e che abbiamo usato per definire l'implicazione.<sup>9</sup>

Dunque le implicazioni con antecedente falso sono vere; osserviamo anche che sono vere le implicazioni

<sup>7</sup>con un piccolo, ma conveniente, abuso di linguaggio si usa chiamare implicazione sia il connettivo ( $\Rightarrow$ ) che una formula come  $\alpha \Rightarrow \beta$  (dove  $\alpha$  e  $\beta$  siano a loro volta formule) in cui il connettivo appare. Similmente accade per gli altri connettivi: può capitare ad esempio di dire che la formula  $\alpha \wedge \beta$  sia una congiunzione

<sup>8</sup>come caso limite, si pensi ad una formula che esprima una frase del tipo "se esiste un polinomio non nullo  $f$  a coefficienti razionali di cui il numero  $\pi$  è radice, allora ...". Accogliendo l'obiezione secondo cui una implicazione il cui antecedente sia (notoriamente) falso non ha senso, dovremmo concludere che questa frase aveva senso verso la metà dell'ottocento ma non lo ha più a partire dal 1882, anno in cui è stato dimostrato (da Ferdinand von Lindemann) che un polinomio come  $f$  non può esistere. O, peggio, non ha più senso per chi conosce il risultato di Lindemann, lo ha per chi non lo conosce. In un linguaggio formale non c'è posto per pasticci del genere.

<sup>9</sup>ad ulteriore conferma, se la frase di partenza fosse stata "per ogni numero intero  $x$  compreso tra 1 e 3 si ha che se  $x > 1$  allora  $x > 2$ ", la frase sarebbe stata falsa. Infatti, le tre proposizioni ottenute sostituendo ad  $x$  i numeri 1, 2, 3 non sono *tutte* vere: quella ottenuta ponendo 2 al posto di  $x$ , cioè " $2 > 1 \Rightarrow 2 > 2$ " è falsa, in accordo col secondo rigo della tavola, avendo l'antecedente vero ed il conseguente falso. Su questi esempi torneremo **più avanti** per chiarirli, si spera, ulteriormente.

con conseguente vero. In effetti, possiamo dire, sinteticamente, che una *implicazione* è vera precisamente quando il suo antecedente è falso o il suo conseguente è vero.

### Esercizi ed Esempi.

**B.1.** Scrivere le tavole di verità di ciascuna delle forme proposizionali: “ $p \wedge (p \vee q)$ ”, “ $(p \wedge q) \wedge r$ ”, “ $(p \wedge q) \vee r$ ”.

**B.2.** Scrivere le tavole di verità di ciascuna delle forme proposizionali: “ $p \Rightarrow (p \vee q)$ ”, “ $(p \wedge q) \Rightarrow r$ ”, “ $(p \wedge q) \Leftrightarrow r$ ”.

**B.3.** Stabilire i valori di verità delle formule e frasi (assumiamo nota la matematica elementare coinvolta): “ $(1 + 1 = 0) \vee (0 + 0 = 0)$ ”; “ $(1 + 1 = 0) \wedge (0 + 0 = 0)$ ”; “ $(1 + 1 = 0) \Rightarrow (0 + 0 = 0)$ ”; “ $\sqrt{2}$  è un numero razionale o un numero irrazionale”; “ $2^5 = 32 \Rightarrow 47 - 1 = 46$ ”.

**B.4.** È molto importante saper ‘tradurre’ espressioni del linguaggio ordinario (della lingua italiana che parliamo quotidianamente) in linguaggio ‘semiformalizzato’, riconoscendo la presenza ed il ruolo dei connettivi proposizionali contenuti nelle frasi. Ad esempio, se indichiamo con  $\alpha$  la frase ‘domani pioverà’ e con  $\beta$  la frase ‘domani prenderò l’ombrellino’, si può rendere con  $\alpha \wedge \beta$  la frase ‘domani pioverà e prenderò l’ombrellino’. Fare lo stesso per le frasi:

- (a) Il supermercato era aperto e non ci sono entrato.
- (b) Il supermercato era aperto ma non ci sono entrato.
- (c) Se vedo Nicola lo saluto.
- (d) Se domenica non piove e vado a Roma,  $2 > 1$ , ma se Marco mangia la pizza allora certamente fioriranno le rose.

**B.5.** Come nell’esercizio precedente, che struttura logica ha la frase: ‘Maria ha cucinato la torta, e Franco non l’ha vista oppure l’ha mangiata’? Scriviamo  $t$  per ‘Maria ha cucinato la torta’,  $v$  per ‘Franco ha visto la torta’ e  $m$  per ‘Franco ha mangiato la torta’. Se facciamo attenzione alla virgola che appare nella frase, concludiamo che questa frase si può rendere con  $t \wedge ((\neg v) \vee m)$ . Bene, come possiamo rendere: ‘Maria ha cucinato la torta e Franco non l’ha vista, oppure l’ha mangiata’? Le due frasi hanno necessariamente gli stessi valori di verità, oppure possono esserci circostanze in cui una è vera e l’altra falsa?

**B.6.** Spiegare la seguente (vecchia e non particolarmente esilarante) storiella: la moglie del logico chiede al marito: ‘Caro, stasera usciamo o restiamo a casa?’’. Il marito risponde: ‘Sì’.

## 3. TAUTOLOGIE

Interrompiamo la lista dei connettivi per introdurre alcune importanti nozioni. Consideriamo una forma proposizionale  $\Phi$  e supponiamo che  $p, q, r, \dots$  siano le variabili proposizionali che *possono* apparire in  $\Phi$  (quest’ultimo fatto si può esprimere scrivendo  $\Phi(p, q, r, \dots)$  per  $\Phi$ ). Se attribuiamo un valore di verità ( $V$  o  $F$ ) a ciascuna delle variabili in  $\Phi$  (in modo consistente, ovviamente: ad esempio sempre lo stesso valore per ogni occorrenza della  $p$ ) possiamo calcolare il valore di verità di  $\Phi$  in funzione di quelli attribuiti a  $p, q, r, \dots$ : è quello che facciamo quando scriviamo una tavola di verità. Si dice che  $\Phi$  è una tautologia se e solo se il valore di verità di  $\Phi$  così calcolato è  $V$ , indipendentemente dai valori attribuiti alle variabili che appaiono in  $\Phi$ . In altri termini,  $\Phi$  è una tautologia se e solo se, nella tavola di verità che la descrive, la colonna intestata da  $\Phi$  contiene esclusivamente  $V$ . Alcuni esempi banali di tautologie sono le forme “ $p \Rightarrow p$ ” e “ $p \Leftrightarrow p$ ”; un altro, che esprime il *principio del terzo escluso*, di cui qualcuno potrebbe avere memoria scolastica, è la forma “ $p \vee (\neg p)$ ”. Per quanto facile, verifichiamo questa tautologia usando una tavola di verità:

terzo escluso		
$p$	$\neg p$	$p \vee (\neg p)$
$V$	$F$	$V$
$F$	$V$	$V$

Come già minacciato, abbiamo qui una tavola di verità in cui la parte destra contiene più di una colonna: la prima contiene i valori di verità della sottoformula  $\neg p$ , usati come passaggio intermedio nel calcolo:

prima si è calcolata questa colonna a partire dalla colonna di  $p$  ed usando la [descrizione di  \$\neg\$](#) , poi si è ottenuta la terza colonna applicando la [disgiunzione](#) alle prime due.

Dualmente, esistono forme proposizionali  $\Phi$  per le quali, calcolando come detto sopra il valore di verità, si ottiene sempre il valore  $F$ . Queste si chiamano *contraddizioni*. Dovrebbe essere chiaro che  $\Phi$  è una contraddizione se e solo se  $\neg\Phi$  è una tautologia. Un famoso esempio di contraddizione (lo si verifichi per esercizio) è la forma  $p \wedge (\neg p)$ , quindi la sua negazione  $\neg(p \wedge (\neg p))$  è una tautologia, il *principio di non contraddizione*. Una forma proposizionale che non sia né una tautologia né una contraddizione si dice *contingente*. Ad esempio, la forma  $p \Rightarrow q$  è contingente perché può assumere, in dipendenza dei valori sostituiti a  $p$  e a  $q$ , sia il valore  $V$  che il valore  $F$ .

Se  $\alpha$  e  $\beta$  sono due forme proposizionali, si dice che  $\alpha$  e  $\beta$  sono *logicamente equivalenti* (o, semplicemente, equivalenti) se e solo se la forma  $\alpha \Leftrightarrow \beta$  è una tautologia. Naturalmente, questa condizione vuol dire che per qualsiasi scelta dei valori di verità attribuiti alle variabili che appaiono o in  $\alpha$  o in  $\beta$ ,  $\alpha$  e  $\beta$  hanno lo stesso valore di verità. O, ancora in altri termini:  $\alpha$  e  $\beta$  sono logicamente equivalenti se e solo se in una tavola di verità in cui appaiano entrambe, ad esse corrisponda la stessa colonna di valori di verità. Vediamo un esempio molto semplice: questa tabella mostra che  $p$  e  $\neg(\neg p)$  sono logicamente equivalenti.

tautologia della doppia negazione

$p$	$\neg p$	$\neg(\neg p)$
$V$	$F$	$V$
$F$	$V$	$F$

Detto diversamente,  $p \Leftrightarrow (\neg(\neg p))$  è una tautologia, quella della *doppia negazione*.

Se una forma  $\alpha$  appare come componente (o *sottoformula*) di una forma proposizionale  $\Phi$  e se  $\beta$  è una forma logicamente equivalente ad  $\alpha$ , è chiaro che sostituendo in  $\Phi$  la sottoformula  $\alpha$  con  $\beta$  si ottiene una forma  $\Phi'$  che sarà logicamente equivalente a  $\Phi$ . Ad esempio, se in  $\Phi$  appare una sottoformula come  $\neg(\neg\varphi)$ , possiamo cancellare le due negazioni e quindi sostituire  $\varphi$  a  $\neg(\neg\varphi)$  ottenendo una forma equivalente a  $\Phi$ . Manipolazioni di questo genere sono spesso utili per calcolare valori di verità senza dover necessariamente far ricorso a tavole di verità.

**Alcune tautologie elementari.** Abbiamo già visto qualche esempio di tautologia. Ne elenchiamo ora altre, in modo più sistematico, spesso attribuendo loro, per nostra comodità, dei nomi che serviranno a poterle richiamare più avanti. La maggior parte di quelle che stiamo per vedere non necessitano di particolari commenti, risultando spesso addirittura ovvie. Chi legge è però caldamente invitato ad esercitarsi verificando che quelle elencate sono effettivamente tautologie. Le prime che incontriamo ricordano, anche nel nome, proprietà abitualmente definite per operazioni algebriche.

idempotenza	commutatività	associatività
$(p \wedge p) \Leftrightarrow p$	$(p \wedge q) \Leftrightarrow (q \wedge p)$	$((p \wedge q) \wedge r) \Leftrightarrow (p \wedge (q \wedge r))$
$(p \vee p) \Leftrightarrow p$	$(p \vee q) \Leftrightarrow (q \vee p)$	$((p \vee q) \vee r) \Leftrightarrow (p \vee (q \vee r))$
	$(p \Leftrightarrow q) \Leftrightarrow (q \Leftrightarrow p)$	$((p \Leftrightarrow q) \Leftrightarrow r) \Leftrightarrow (p \Leftrightarrow (q \Leftrightarrow r))$

Sull'associatività di  $\wedge$  e  $\vee$ , osserviamo che  $(p \wedge q) \wedge r$  risulta vera se e solo se sono contemporaneamente vere sia  $p$  che  $q$  che  $r$  (lo stesso vale per  $p \wedge (q \wedge r)$ , altrimenti non avremmo una tautologia), mentre  $(p \vee q) \vee r$  (ovvero,  $p \vee (q \vee r)$ ) è vera se e solo se è vera almeno una tra  $p$ ,  $q$  ed  $r$ . Più in generale (come accade in algebra), a partire da queste tautologie è possibile (ma noioso) provare che, qualunque sia l'intero positivo  $k$  le forme proposizionali in cui appaiano tutte e sole le variabili  $p_1, p_2, \dots, p_k$ , delle parentesi e, tra i connettivi, solo  $\wedge$  sono equivalenti tra loro, indipendentemente dall'ordine in cui appaiano le variabili, dalle eventuali ripetizioni e dal modo in cui esse sono raggruppate (cioè da come sono disposte le parentesi).<sup>10</sup> Per queste forme si può allora rinunciare all'uso delle parentesi e scrivere semplicemente  $p_1 \wedge p_2 \wedge p_3 \wedge \dots \wedge p_k$  (omettendo, magari, anche le ripetizioni) o  $\bigwedge_{i=1}^k p_i$  per indicare una qualunque di queste forme. Questa forma assume il valore  $V$  se e solo se valgono  $V$  tutte le  $p_i$ . Simile

<sup>10</sup>ad esempio, per  $k = 4$ , le forme  $(p_1 \wedge p_2) \wedge (p_3 \wedge p_4)$ ,  $p_1 \wedge ((p_2 \wedge p_3) \wedge p_4)$ ,  $((p_2 \wedge p_4) \wedge (p_1 \wedge p_3)) \wedge p_4$  ed infinite altre sono tra loro equivalenti

enunciato vale per  $\vee$  e giustifica l'uso di scritture come  $p_1 \vee p_2 \vee p_3 \vee \dots \vee p_k$  o  $\bigvee_{i=1}^k p_i$ , questa forma vale  $V$  se e solo se almeno una tra le  $p_i$  vale  $V$ .

Anche a proposito dell'associatività di  $\Leftrightarrow$  qualche commento può essere utile. La verifica del fatto che tratti di una tautologia, cioè del fatto che  $(p \Leftrightarrow q) \Leftrightarrow r$  e  $p \Leftrightarrow (q \Leftrightarrow r)$  sono logicamente equivalenti, è contenuta nella tavola di verità (la cui verifica è lasciata a chi legge):

associatività della equivalenza					
$p$	$q$	$r$	$(p \Leftrightarrow q) \Leftrightarrow r$	$p \Leftrightarrow (q \Leftrightarrow r)$	
$V$	$V$	$V$	$V$	$V$	
$V$	$V$	$F$	$F$	$F$	
$V$	$F$	$V$	$F$	$F$	
$V$	$F$	$F$	$V$	$V$	
$F$	$V$	$V$	$F$	$F$	
$F$	$V$	$F$	$V$	$V$	
$F$	$F$	$V$	$V$	$V$	
$F$	$F$	$F$	$F$	$F$	

Si noti che  $(p \Leftrightarrow q) \Leftrightarrow r$  vale vera se e solo se esattamente uno o tutti e tre tra  $p$ ,  $q$  e  $r$  valgono vero.

*Distributività.* Altre tautologie che ricordano da vicino proprietà algebriche sono le *leggi distributive* (di  $\wedge$  rispetto a  $\vee$  e viceversa):

$$\begin{aligned} p \wedge (q \vee r) &\iff (p \wedge q) \vee (p \wedge r) \\ p \vee (q \wedge r) &\iff (p \vee q) \wedge (p \vee r), \end{aligned}$$

Verifichiamo la prima delle due utilizzando una tavola di verità, questa volta con tutti i passaggi intermedi:

$p$	$q$	$r$	$q \vee r$	$p \wedge (q \vee r)$	$p \wedge q$	$p \wedge r$	$(p \wedge q) \vee (p \wedge r)$
$V$	$V$	$V$	$V$	$V$	$V$	$V$	$V$
$V$	$V$	$F$	$V$	$V$	$V$	$F$	$V$
$V$	$F$	$V$	$V$	$V$	$F$	$V$	$V$
$V$	$F$	$F$	$F$	$F$	$F$	$F$	$F$
$F$	$V$	$V$	$V$	$F$	$F$	$F$	$F$
$F$	$V$	$F$	$V$	$F$	$F$	$F$	$F$
$F$	$F$	$V$	$V$	$F$	$F$	$F$	$F$
$F$	$F$	$F$	$F$	$F$	$F$	$F$	$F$

Confrontando la quinta e l'ottava colonna, che coincidono in tutto tranne che nell'intestazione, si ottiene il risultato.

### Esercizi.

**C.1.** Verificare la seconda delle tautologie appena enunciate (cioè la distributività della disgiunzione rispetto alla congiunzione).

**C.2.** La forma proposizionale  $(p \wedge q) \Rightarrow (p \vee q)$  è una tautologia, mentre  $(p \vee q) \Rightarrow (p \wedge q)$  non lo è. Scrivere la tavola di verità di quest'ultima. È possibile scrivere una forma proposizionale più breve (nel senso ovvio) di  $(p \vee q) \Rightarrow (p \wedge q)$  e che sia equivalente a questa?

**C.3.** Verificare la tautologia  $(p \Rightarrow (q \Rightarrow r)) \iff ((p \Rightarrow q) \Rightarrow (p \Rightarrow r))$  (distributività da sinistra della implicazione rispetto a se stessa).

*Leggi di De Morgan.* Quando è che una proposizione della forma  $p \wedge q$  è falsa? Risposta: quando (e solo quando) è falsa *almeno una* tra  $p$  e  $q$ . Questo è evidente dalla [tavola di verità che descrive la congiunzione](#). La risposta che qualche volta capita di ricevere: ‘quando sia  $p$  che  $q$  sono false’ è sbagliata: la proposizione è sicuramente falsa in questo caso, ma lo è anche in altri. Ad esempio, perché non sia vero che io abbia preso il treno e sia arrivato a Firenze non è necessario che io non abbia preso il treno e non sia arrivato a Firenze, la frase è falsa anche se io ho preso il treno e mi sono fermato a Roma, ad esempio, o anche se io sono andato a Firenze, ma in auto.

Dualmente, una proposizione della forma  $p \vee q$  è falsa precisamente quando *sia p che q* sono false (come mostra la [tavola di verità della disgiunzione](#)), quindi la negazione di ‘prendo l’auto o vado a piedi’ è ‘non prendo l’auto e non vado a piedi’. Tutto questo è espresso da due tautologie molto importanti, note come *leggi di De Morgan*:

$$\begin{aligned} (\neg(p \wedge q)) &\iff ((\neg p) \vee (\neg q)) \\ (\neg(p \vee q)) &\iff ((\neg p) \wedge (\neg q)) \end{aligned} \quad (\text{De Morgan})$$

Dunque, una disgiunzione (o una congiunzione) si negano negando i due termini che stiamo disgiungendo (o congiungendo) e, contemporaneamente, scambiando tra loro i simboli  $\wedge$  e  $\vee$ .

### Esercizi.

**D.1.** Negare ciascuna delle frasi: “Mario corre e Maria nuota”; “La bottiglia è vuota oppure tappata”.

**D.2.** Negare la frase: ‘Alice ha i capelli biondi ricci’. (Si chiede che anche la negazione inizi con ‘Alice ha i capelli ...’). Attenzione: quale connettivo proposizionale è nascosto nella frase?)

**D.3.** Usando le leggi di De Morgan, negare  $p \vee (\neg(q \wedge (\neg p)))$ . Ciò che si chiede è scrivere una formula che sia equivalente alla negazione di quella data e che non abbia  $\neg$  come primo simbolo.

**D.4.** Come per l’esercizio precedente, negare ciascuna delle due formule: “ $p \wedge q \wedge r$ ” e “ $(p \vee q) \wedge ((p \vee r) \wedge (q \vee s))$ ”.

*Tautologie sulla implicazione.* Le tautologie sulla implicazione hanno grande importanza: il carattere meno intuitivo di questo connettivo le rende meno ovvie di quelle che abbiamo incontrato finora, ma l’uso frequentissimo che si fa in matematica del connettivo di implicazione rende queste tautologie uno strumento utilissimo. Vale dunque la pena di soffermarsi con una certa attenzione su di esse.

In primo luogo, il connettivo “ $\Rightarrow$ ”, a differenza degli altri connettivi binari analizzati finora, non è **commutativo**; vale a dire: le forme “ $p \Rightarrow q$ ” e “ $q \Rightarrow p$ ” non sono equivalenti tra loro. Basta una tavola di verità per convincersene:

$p$	$q$	$p \Rightarrow q$	$q \Rightarrow p$
$V$	$V$	$V$	$V$
$V$	$F$	$F$	$V$
$F$	$V$	$V$	$F$
$F$	$F$	$V$	$V$

Prendiamo nota del fatto che spesso si scrive “ $p \Leftarrow q$ ” per “ $q \Rightarrow p$ ”. Si può considerare questo simbolo “ $\Leftarrow$ ” come un ulteriore connettivo binario (*implicazione inversa*), definito appunto dall’essere  $p \Leftarrow q$  logicamente equivalente a  $q \Rightarrow p$ .

Come chi legge queste note certamente già sa, la congiunzione di una implicazione e della corrispondente implicazione inversa equivale alla doppia implicazione, cioè vale la tautologia:

$$(p \Leftarrow q) \iff ((p \Rightarrow q) \wedge (q \Rightarrow p)), \quad (\text{tautologia della doppia implicazione})$$

come mostrato dalla tavola di verità:

$p$	$q$	$p \Leftrightarrow q$	$p \Rightarrow q$	$p \Leftarrow q$	$(p \Rightarrow q) \wedge (p \Leftarrow q)$
$V$	$V$	$V$	$V$	$V$	$V$
$V$	$F$	$F$	$F$	$V$	$F$
$F$	$V$	$F$	$V$	$F$	$F$
$F$	$F$	$V$	$V$	$V$	$V$

Non basteranno mai gli avvertimenti e le preghiere rivolte agli studenti perché facciano attenzione a non confondere tra loro implicazione, implicazione inversa ed equivalenza. Si tratta, lo abbiamo visto in dettaglio, di connettivi che hanno funzioni logiche ben diverse; confonderli porta quasi certamente ad errori di ragionamento, spesso molto gravi.<sup>11</sup>

Forse questo è un punto adatto per elencare alcune delle tante espressioni che vengono utilizzate in matematica per rendere nel linguaggio ordinario i connettivi di implicazione, implicazione inversa ed equivalenza. In ciascuna colonna si possono leggere frasi che traducono la formula nell'intestazione:

$p \Rightarrow q$	$p \Leftarrow q$	$p \Leftrightarrow q$
Se $p$ allora $q$		
$p$ solo se $q$	$p$ se $q$	$p$ se e solo se $q$
$p$ è condizione sufficiente per $q$	$p$ è condizione necessaria per $q$	$p$ è condizione necessaria e sufficiente per $q$

Come si è visto, dunque, il connettivo di equivalenza può essere ridotto a quelli di implicazione e congiunzione. Anche il connettivo di implicazione può essere espresso in termini di altri connettivi (disgiunzione e negazione). Si ha infatti questa tautologia:

$$(p \Rightarrow q) \iff ((\neg p) \vee q), \quad (\text{implicazione come disgiunzione})$$

che segue direttamente dalle tavole di verità di [implicazione](#) e [disgiunzione](#) e che avevamo sostanzialmente già osservato nelle [ultime righe della sezione in cui è stata introdotta l'implicazione](#), dove abbiamo notato che una implicazione è vera se e solo se il suo antecedente è falso o il suo conseguente è vero. Da questa tautologia se ne può facilmente dedurre un'altra, la *legge di contrapposizione*:

$$((p \Rightarrow q)) \iff ((\neg q) \Rightarrow (\neg p)). \quad (\text{legge di contrapposizione})$$

Il passaggio è il seguente:  $p \Rightarrow q$  equivale a  $(\neg p) \vee q$ , ovvero a  $q \vee (\neg p)$  (per la [commutatività di  \$\vee\$](#) ), ovvero a  $(\neg(\neg q)) \vee (\neg p)$  (per la tautologia della [doppia negazione](#)), ma quest'ultima, per la tautologia dell'[implicazione come disgiunzione](#), equivale a  $(\neg q) \Rightarrow (\neg p)$ .

Molto importante, ed anch'essa immediata dalla [tavola di verità della implicazione](#), è la tautologia che mostra come negare una implicazione: questa è falsa precisamente quando l'antecedente è vero ed il conseguente è falso.

$$(\neg(p \Rightarrow q)) \iff (p \wedge (\neg q)). \quad (\text{negazione dell'implicazione})$$

Un'altra tautologia di uso frequentissimo è quella della *transitività dell'implicazione*:

$$((p \Rightarrow q) \wedge (q \Rightarrow r)) \Rightarrow (p \Rightarrow r). \quad (\text{transitività dell'implicazione})$$

Piuttosto che scrivere una (noiosa e lunga) tavola di verità, verifichiamo questa tautologia utilizzando un metodo che è spesso molto conveniente quando si ha a che fare con le implicazioni. Per provare che la formula risulti vera, indipendentemente dal valore di verità attribuito a  $p$ ,  $q$ ,  $r$ , poveremo che in nessun caso essa può risultare falsa. Perché la formula sia falsa occorre che sia vero l'antecedente  $((p \Rightarrow q) \wedge (q \Rightarrow r))$  e falso il conseguente  $(p \Rightarrow r)$ . La prima condizione significa che sono vere  $p \Rightarrow q$  e  $q \Rightarrow r$ , la seconda che sia vera  $p$  e falsa  $r$ . Ora, assumendo queste condizioni, sono in particolare vere  $p$  e  $p \Rightarrow q$ ; da ciò segue subito che  $q$  è vera (se  $p$  è vera ma  $q$  è falsa, allora  $p \Rightarrow q$  è falsa!). Quindi, se la nostra formula è falsa, risultano vere  $p$  e  $q$ , ma falsa  $r$ . Tuttavia, in questo caso,  $q \Rightarrow r$  è falsa, mentre avevamo detto che, perché la formula sia falsa,  $q \Rightarrow r$  deve essere vera. Questo ragionamento mostra che

<sup>11</sup>come quello del contadino che, sapendo che se piove si esce con l'ombrelllo, pensa che sia sufficiente prendere l'ombrelllo per garantire un po' di pioggia ai suoi campi.

la formula considerata, cioè  $((p \Rightarrow q) \wedge (q \Rightarrow r)) \Rightarrow (p \Rightarrow r)$ , non può essere falsa in nessun caso, quindi è una tautologia.

L'idea esemplificata da questa dimostrazione consiste in questo: imporre che una implicazione sia falsa fornisce immediatamente due informazioni: i valori di verità di antecedente e conseguente. Dunque può essere conveniente, nello studiare una implicazione, analizzare subito le conseguenze dell'ipotesi che essa sia falsa.

È chiaro che dalla **transitività dell'implicazione** (e dalla **tautologia della doppia implicazione**) si possono dedurre molte altre tautologie che coinvolgano i connettivi  $\Rightarrow$ ,  $\Leftarrow$  e  $\Leftrightarrow$ , come ad esempio la **transitività della equivalenza**:  $((p \Leftrightarrow q) \wedge (q \Leftrightarrow r)) \Rightarrow (p \Leftrightarrow r)$  o altre come  $((p \Leftrightarrow q) \wedge (q \Rightarrow r) \wedge (r \Rightarrow t)) \Rightarrow (p \Rightarrow t)$ .

Concludiamo accennando a una notazione abbreviata, molto intuitiva e di uso comune, quella delle cosiddette catene di implicazioni, che esprimono congiunzioni tra formule che sono a loro volta implicazioni, implicazioni inverse o equivalenze. Dovrebbe bastare qualche esempio: espressioni come  $\alpha \Rightarrow \beta \Leftrightarrow \gamma$  oppure  $\alpha \Leftarrow \beta \Rightarrow \gamma \Rightarrow \delta$  (senza parentesi), vengono usate al posto di  $(\alpha \Rightarrow \beta) \wedge (\beta \Leftrightarrow \gamma)$  e  $(\alpha \Leftarrow \beta) \wedge (\beta \Rightarrow \gamma) \wedge (\gamma \Rightarrow \delta)$  rispettivamente. Dunque, la transitività della implicazione si può riscrivere come  $(p \Rightarrow q \Rightarrow r) \Rightarrow (p \Rightarrow r)$ .

### Esercizi ed Osservazioni.

**E.1.** Torniamo sull'[Esercizio B.5](#). La prima frase proposta in quell'esercizio si potrebbe rendere, in modo equivalente, con  $t \wedge (v \Rightarrow m)$ ? Oppure con  $t \wedge (m \Rightarrow v)$ ?

**E.2.** Studiare la forma  $(p \Rightarrow q) \vee (p \Leftarrow q)$ . Confrontarla col secondo membro (il termine a destra di " $\Leftrightarrow$ ") nella **tautologia della doppia implicazione**.

**E.3.** Utilizzando le leggi di De Morgan, si determini la negazione di  $(\neg p) \vee q$ . Si confronti quanto ottenuto con le tautologie della **implicazione come disgiunzione** e con quella della **negazione dell'implicazione**.

**E.4.** La **legge di contrapposizione** è alla base di una tecnica dimostrativa di uso molto frequente. Per dimostrare una tesi  $T$  partire da una ipotesi  $H$ , cioè per dimostrare l'implicazione  $H \Rightarrow T$ , si può assumere falsa  $T$  e dedurre da questa assunzione la falsità di  $H$ . In altri termini, ciò che si fa in questo modo è dimostrare l'implicazione  $(\neg T) \Rightarrow (\neg H)$ . Per la legge di contrapposizione, questa formula equivale ad  $H \Rightarrow T$ , quella che si voleva dimostrare. Dimostrazioni condotte in questo modo si chiamano *dimostrazioni per contrapposizione*.

**E.5.** Si provi che le forme proposizionali " $(p \wedge q) \Rightarrow r$ ", " $p \Rightarrow (q \Rightarrow r)$ " e " $q \Rightarrow (p \Rightarrow r)$ " sono tra loro logicamente equivalenti. Si consiglia di non utilizzare, a questo scopo, tavole di verità ma di ragionare come fatto per la **transitività dell'implicazione**. Questo consiglio si estende anche agli esercizi che seguono.

**E.6.** Verificare la tautologia " $(p \wedge (p \Rightarrow q)) \Rightarrow q$ ", nota come *legge dell'inferenza*.

**E.7.** Tornare all'[Esercizio C.3](#) (se lo si è svolto) e ripetere (o farla ex-novo) la verifica senza usare tavole di verità. La forma " $((p \Rightarrow q) \Rightarrow r) \Leftrightarrow ((p \Rightarrow r) \Rightarrow (q \Rightarrow r))$ " è una tautologia?

**E.8.** Verificare le tautologie " $(p \Rightarrow (q \wedge r)) \Leftrightarrow ((p \Rightarrow q) \wedge (p \Rightarrow r))$ " e " $(p \Rightarrow (q \vee r)) \Leftrightarrow ((p \Rightarrow q) \vee (p \Rightarrow r))$ " (distributività da sinistra di  $\Rightarrow$  rispetto a  $\wedge$  e a  $\vee$ ). Suggerimento: si usi la tautologia della **implicazione come disgiunzione** per trasformare tutte le implicazioni in disgiunzioni che coinvolgano  $\neg p$ .

Le forme " $((q \wedge r) \Rightarrow p) \Leftrightarrow ((q \Rightarrow p) \wedge (r \Rightarrow p))$ " e " $((q \vee r) \Rightarrow p) \Leftrightarrow ((q \Rightarrow p) \vee (r \Rightarrow p))$ " sono tautologie?

## 4. ALTRI CONNETTIVI BINARI

Per completare la nostra discussione sul calcolo proposizionale vanno ancora menzionati alcuni connettivi binari, che hanno interesse di per sé e sono spesso utilizzati in informatica. Il più importante tra questi è la *disgiunzione esclusiva* ( $\dot{\vee}$ , o XOR), a cui abbiamo già fatto [cenno](#). Altri due sono le negazioni della congiunzione e della disgiunzione, rispettivamente NAND e NOR (da leggere come not-and e not-or), che sono anche noti, rispettivamente, come operazione di Sheffer (o *stroke*, indicata con  $|$  o  $\uparrow$ ) e operazione di Pierce (e indicata con  $\downarrow$ ). Le tavole di verità sono:

$p$	$q$	$p \dot{\vee} q$	$p$	$q$	$p \mid q$	$p$	$q$	$p \downarrow q$
		$p \text{ XOR } q$			$p \text{ NAND } q$			$p \text{ NOR } q$
$V$	$V$	$F$	$V$	$V$	$F$	$V$	$V$	$F$
$V$	$F$	$V$	$V$	$F$	$V$	$F$	$F$	$F$
$F$	$V$	$V$	$F$	$V$	$V$	$F$	$V$	$F$
$F$	$F$	$F$	$F$	$F$	$V$	$F$	$F$	$V$

È chiaramente giustificato quanto abbiamo appena detto su NAND e NOR: questi connettivi negano  $\wedge$  e  $\vee$ , nel senso che “ $(p \text{ NAND } q) \Leftrightarrow (\neg(p \wedge q))$ ” e “ $(p \text{ NOR } q) \Leftrightarrow (\neg(p \vee q))$ ” sono tautologie. Allo stesso modo, XOR nega  $\Leftrightarrow$ . Si ha infatti una utilissima serie di tautologie, che si possono esprimere come catena di equivalenze:

$$(\neg(p \Leftrightarrow q)) \Leftrightarrow ((\neg p) \Leftrightarrow q) \Leftrightarrow (p \Leftrightarrow (\neg q)) \Leftrightarrow (p \text{ XOR } q). \quad (\text{negazione dell'equivalenza})$$

Ricordiamo cosa significa: le quattro forme proposizionali elencate sono a due a due logicamente equivalenti. Che lo siano si vede subito: ciascuna di esse è vera quando e solo quando  $p$  e  $q$  hanno diversi valori di verità, cioè una vale ‘vero’, l’altra vale ‘falso’. Notiamo che queste tautologie ci mostrano come possiamo importare ed esportare (portare ‘dentro’ o ‘fuori’ a nostro piacimento) il simbolo di negazione da una equivalenza a ciascuno dei termini che vi appaiono. Usando questo fatto possiamo dimostrare una proprietà molto importante: l’associatività del connettivo XOR.<sup>12</sup>

Consideriamo infatti la forma  $p \text{ XOR } (p \text{ XOR } r)$ . La catena di equivalenze appena osservata mostra che  $p \text{ XOR } (q \text{ XOR } r)$  equivale a  $\neg(p \Leftrightarrow (q \text{ XOR } r))$  e questa (importando la negazione al secondo termine dell’equivalenza), equivale a  $p \Leftrightarrow (\neg(q \text{ XOR } r))$ , cioè a  $p \Leftrightarrow (q \Leftrightarrow r)$ . Abbiamo dunque la tautologia

$$(p \text{ XOR } (q \text{ XOR } r)) \Leftrightarrow (p \Leftrightarrow (q \Leftrightarrow r)).$$

Allo stesso modo (oppure usando quest’ultima tautologia insieme alla *commutatività*, ovvia, di XOR<sup>13</sup>) e a quella di  $\Leftrightarrow$  si verifica la tautologia

$$((p \text{ XOR } q) \text{ XOR } r) \Leftrightarrow ((p \Leftrightarrow q) \Leftrightarrow r).$$

Da queste due, e dall’*associatività di  $\Leftrightarrow$*  si ricava la tautologia che volevamo provare:

$$(p \text{ XOR } (q \text{ XOR } r)) \Leftrightarrow ((p \text{ XOR } q) \text{ XOR } r). \quad (\text{associatività di XOR})$$

Altre due facili tautologie che riguardano XOR sono espresse in questa catena di equivalenze:

$$(p \text{ XOR } q) \Leftrightarrow ((p \wedge (\neg q)) \vee (q \wedge (\neg p))) \Leftrightarrow ((p \vee q) \wedge (\neg(p \wedge q))). \quad (\text{esplicitazione di XOR})$$

Anche queste equivalenze si provano facilmente osservando che, evidentemente, sia  $((p \wedge (\neg q)) \vee (q \wedge (\neg p)))$  che  $((p \vee q) \wedge (\neg(p \wedge q)))$  sono vere se e solo esattamente uno tra  $p$  e  $q$  è vero, ma si veda a questo riguardo anche l’Esercizio F.1. È poi importante (per lo studio delle strutture booleane) la tautologia

$$(p \wedge (q \text{ XOR } r)) \Leftrightarrow ((p \wedge q) \text{ XOR } (p \wedge r)), \quad (\text{distributività di } \wedge \text{ rispetto a XOR})$$

di verifica diretta.

L’ultima osservazione che facciamo prima di chiudere con il calcolo proposizionale evidenzia una particolarità che rende interessanti i connettivi NAND e NOR. Sappiamo che la nostra lista di connettivi è piuttosto ridondante. Non avremmo avuto bisogno, ad esempio, di definire in modo indipendente il connettivo di equivalenza, ma avremmo potuto considerare la scrittura “ $p \Leftrightarrow q$ ” come una semplice abbreviazione di “ $(p \Rightarrow q) \wedge (q \Rightarrow p)$ ”, dal momento che la *tautologia della doppia implicazione* ci dice che le due formule sono equivalenti. Similmente, la *tautologia della implicazione come disgiunzione* mostra come sia possibile fare a meno anche del connettivo  $\Rightarrow$  ed esprimere tutte le implicazioni (e le implicazioni inverse) utilizzando  $\neg$  e  $\vee$ ; in modo ancora più immediato NAND e NOR si ottengono da  $\neg$ ,  $\vee$  e  $\wedge$ , che

<sup>12</sup>il fatto che XOR sia associativo permette, tra l’altro, di definire in modo sintetico le strutture booleane, che hanno enorme importanza nell’informatica teorica.

<sup>13</sup>cioè la tautologia  $(p \text{ XOR } q) \Leftrightarrow (q \text{ XOR } p)$ . Anche NOR e NAND sono commutative; tutto ciò è evidente dalle *descrizioni* che abbiamo dato di questi connettivi.

certamente bastano per esprimere anche XOR. Dunque possiamo esprimere tutto il calcolo proposizionale usando solo i connettivi  $\neg$ ,  $\vee$  e  $\wedge$ . Meglio ancora: dalle leggi di De Morgan (e dalla tautologia della doppia negazione) seguono le tautologie

$$(p \vee q) \Leftrightarrow (\neg(\neg p) \wedge (\neg q)) \quad \text{e} \quad (p \wedge q) \Leftrightarrow (\neg(\neg p) \vee (\neg q)),$$

che mostrano come la disgiunzione si possa esprimere utilizzando  $\neg$  e  $\wedge$ , mentre la congiunzione si può esprimere utilizzando  $\neg$  e  $\vee$ . Dunque tutte le formule del calcolo proposizionale possono essere espresse (a meno di equivalenze) usando solo due connettivi:  $\neg$  ed uno a scelta tra  $\wedge$  e  $\vee$ . Se siamo davvero interessati a ridurre il numero dei connettivi usati, si può addirittura fare di meglio: basta un solo connettivo, uno a scelta tra NAND e NOR. Abbiamo infatti due tautologie (di semplicissima verifica):

$$(p \text{ NOR } p) \Leftrightarrow (\neg p) \Leftrightarrow (p \text{ NAND } p)$$

che garantiscono come l'uso della negazione possa sempre essere sostituito dall'uso di una qualsiasi tra NAND e NOR. Dunque, il connettivo NAND permette di esprimere  $\neg$ , e quindi  $\wedge$  (perché  $p \wedge q$  equivale a  $\neg(p \text{ NAND } q)$ , cioè a  $(p \text{ NAND } q) \text{ NAND } (p \text{ NAND } q)$ ). Per quanto visto sopra, possiamo concludere che *ogni formula del calcolo proposizionale equivale ad una formula in cui l'unico connettivo che appaia è NAND*. Allo stesso modo, dal momento che NOR basta per esprimere la negazione e quindi la disgiunzione (tramite la tautologia  $(p \vee q) \Leftrightarrow (\neg(p \text{ NOR } q))$ , *anche NOR ha la stessa proprietà*). È questo il motivo che rende le porte NAND e le porte NOR così utili per la progettazione di circuiti elettronici.

### Esercizi.

**F.1.** Usando le leggi distributive e le leggi di De Morgan, verificare direttamente che la forma proposizionale  $(p \wedge (\neg q)) \vee (q \wedge (\neg p))$  è equivalente a  $(p \vee q) \wedge ((\neg q) \vee (\neg p))$  e quindi a  $(p \vee q) \wedge (\neg(p \wedge q))$ . Utilizzando le tautologie della doppia implicazione e della implicazione come disgiunzione verificare poi che anche  $(\neg p) \Leftrightarrow q$  è equivalente a  $(p \vee q) \wedge ((\neg q) \vee (\neg p))$ . Questo fornisce una dimostrazione alternativa per le due tautologie che abbiamo chiamato **esplicitazione di XOR**.

**F.2.** Verificare in dettaglio la **distributività della congiunzione rispetto a XOR**. Basta usare una tavola di verità, ma si può ragionare in modo più sintetico così: se  $p$  vale ‘falso’ entrambi i membri della equivalenza sono falsi, se  $p$  vale ‘vero’, entrambi sono equivalenti a  $q$  XOR  $r$ . Completare il ragionamento verificando tutti i passaggi.

**F.3.** Vale la tautologia “ $(p \vee (q \text{ XOR } r)) \Leftrightarrow ((p \vee q) \text{ XOR } (p \vee r))$ ” (distributività della disgiunzione rispetto a XOR)?

**F.4.** Scrivere una forma proposizionale equivalente a  $p \Rightarrow q$  in cui appaiano solo le variabili  $p$  e  $q$ , il connettivo NAND e, eventualmente, parentesi.

## 5. QUANTIFICATORI

Consideriamo la formula “ $x > 1$ ” del linguaggio dell’aritmetica che abbiamo già (informalmente) introdotto nella prima sezione di queste note. In accordo con quanto scritto lì, questa formula non ha un valore di verità, perché non è una proposizione. Abbiamo però a disposizione una idea intuitiva di ‘sostituzione’<sup>14</sup> che ci permette di estendere la nostra nozione di verità, valutando la formula per ciascuno dei numeri che possono essere sostituiti alla variabile  $x$ .<sup>15</sup> In termini semplici, non abbiamo difficoltà a dire che la formula è vera “per  $x = 10$ ” (cioè sostituendo ad  $x$  il numero 10) ed è falsa per  $x = 0$ . Se chiamiamo  $\varphi$ , o  $\varphi(x)$ , la nostra formula, possiamo indicare con  $\varphi(10)$  e  $\varphi(0)$ , rispettivamente, le formule “ottenute da  $\varphi$  sostituendo ad  $x$  i numeri 10 e 0”, nell’ordine, quindi  $\varphi(10)$  è la formula (chiusa! e vera)  $10 > 1$  mentre  $\varphi(0)$  è la formula (chiusa e falsa)  $0 > 1$ .

Può capitare che una formula risulti vera per ogni possibile sostituzione delle variabili. In questo caso diremo che la formula è *valida*.<sup>16</sup> Ad esempio, sono valide le formule  $x = x$  o anche le formule ricavate

<sup>14</sup>Idea che naturalmente, ma con qualche fatica, si potrebbe formalizzare. A qualcosa in più accenneremo **tra poche pagine**, dopo aver discusso di variabili libere e vincolate

<sup>15</sup>Scriviamo ‘numeri’ perché il linguaggio che stiamo usando è quello dell’aritmetica e quindi implicitamente assumiamo che le variabili possano indicare solo oggetti che chiamiamo numeri interi, qualsiasi cosa essi siano.

<sup>16</sup>Ancora una volta dobbiamo sottolineare che stiamo fornendo una trattazione molto semplificata della materia. Quella di validità di una formula è in realtà una nozione più ricca di quanto risulta qui, ma la sua definizione completa dipende da concetti che non ci è necessario toccare e non toccheremo.

da tautologie, come ad esempio  $\varphi \vee (\neg\varphi)$ , per qualsiasi formula  $\varphi$ , che si ottiene rimpiazzando  $p$  con  $\varphi$  nella tautologia [principio del terzo escluso](#). Useremo espressioni come ‘vale l’implicazione  $\varphi \Rightarrow \psi$ ’ per dire che la formula  $\varphi \Rightarrow \psi$  è valida, oppure ‘ $\varphi$  e  $\psi$  sono equivalenti’ per dire che  $\varphi \Leftrightarrow \psi$  è valida; questo per *arbitrarie* (cioè non necessariamente chiuse) formule  $\varphi$  e  $\psi$ .

Torniamo al nostro discorso. La nozione di sostituzione ci permette di introdurre due nuovi simboli logici, che svolgono un ruolo centrale in un capitolo della logica chiamato *calcolo dei predicati*. Questi simboli sono il *quantificatore universale*  $\forall$  ed il *quantificatore esistenziale*  $\exists$ .<sup>17</sup> La sintassi è semplice: se  $\varphi$  è una formula ed  $x$  è una variabile allora anche “ $\forall x(\varphi)$ ” e “ $\exists x(\varphi)$ ” sono formule; per maggior leggibilità si può anche scrivere “ $(\forall x)(\varphi)$ ” e “ $(\exists x)(\varphi)$ ” o, come al solito, scrivere  $\varphi(x)$  al posto di  $\varphi$  per evidenziare la possibilità che  $x$  appaia in  $\varphi$ .

La prima formula,  $\forall x(\varphi)$ , è quella che si chiama una *formula universale*<sup>18</sup> e si legge ‘per ogni  $x$ ,  $\varphi$ '. Questa formula esprime la contemporanea affermazione di tutte le formule  $\varphi(a)$  ottenute sostituendo ad  $x$  ogni possibile valore  $a$ . Può aiutare pensare a questa formula come ad una versione generalizzata della congiunzione: la congiunzione tra tutte le formule  $\varphi(a)$  ottenute da  $\varphi$  per sostituzione di  $x$  con  $a$ , per ogni possibile scelta di  $a$ .<sup>19</sup> Nel caso in cui ciascuna delle formule  $\varphi(a)$  così ottenute sia chiusa (cioè una proposizione), la formula  $\forall x(\varphi)$  esprime il fatto che ciascuna delle  $\varphi(a)$  è vera. Ad esempio, usando come  $\varphi$  la formula  $x = x$  otteniamo la formula  $\forall x(x = x)$ , che è una proposizione (vedremo come mai) vera perché qualsiasi sia l’oggetto che sostituiamo ad  $x$ , questo oggetto è uguale a se stesso. Invece, se partiamo dalla formula  $x > 1$  dell’aritmetica usata come esempio nel paragrafo precedente, otteniamo  $\forall x(x > 1)$ , che è ancora una proposizione ma è falsa, perché non tutti i numeri interi verificano la condizione di essere maggiori di uno.

Se le formule universali possono essere pensate come una sorta di congiunzione generalizzata, le *formule esistenziali*, cioè quelle introdotte dal quantificatore  $\exists$ , possono invece essere pensate come disgiunzioni generalizzate. Se  $x$  è una variabile e  $\varphi = \varphi(x)$  una formula,  $\exists x(\varphi)$  (che si può leggere: ‘esiste un  $x$  tale che  $\varphi$ ’) esprime l’affermazione di *almeno una* tra le formule  $\varphi(a)$  ottenute sostituendo ad  $x$  ogni possibile valore  $a$ . Ad esempio, nel linguaggio dell’aritmetica la formula  $\exists x(x > 1)$  è una proposizione (lo vedremo) ed è vera perché, ad esempio, è vera la formula “ $3 > 1$ ” ottenuta sostituendo 3 ad  $x$  in “ $x > 1$ ”. Invece la formula  $\exists x(x \neq x)$  è falsa.

Oltre a  $\forall$  ed  $\exists$  esistono altri quantificatori. Quello di uso più frequente è “ $\exists!$ ”. Se  $\varphi$  è una formula ed  $x$  è una variabile, la formula “ $\exists!x(\varphi)$ ” si legge “esiste uno ed un solo  $x$  tale che  $\varphi$ ” ed afferma  $\varphi(a)$  per uno dei possibili valori  $a$  che possono essere sostituiti ad  $x$ , negando  $\varphi(b)$  per ogni  $b$  diverso da  $a$ ; in termini più semplici:  $\varphi$  è verificata da  $a$  e solo da  $a$ . In modo sintetico e più formale, se  $y$  è una variabile (diversa da  $x$ ) che non appare in  $\varphi$ , questo quantificatore è definito dall’equivalenza:<sup>20</sup>

$$\exists!x(\varphi(x)) \iff \exists x(\forall y(\varphi(y) \Leftrightarrow y = x)). \quad (\text{descrizione di } \exists!)$$

È un utile esercizio comprendere questa equivalenza, cioè il fatto che il suo membro a destra (vale a dire  $\exists x(\forall y(\varphi(y) \Leftrightarrow y = x))$ ) esprime proprio ciò che vogliamo esprimere il membro a sinistra ( $\exists!x(\varphi(x))$ ). Commenti a questo proposito sono nell’[Osservazione G.1](#).

**Variabili libere e vincolate.** In una formula come  $\forall x(\varphi)$  o come  $\exists x(\varphi)$  (anche se queste appaiono come sottoformule di formule più complesse) si dice che le occorrenze di  $x$  sono *vincolate* (dal quantificatore  $\forall$  o dal quantificatore  $\exists$ ). Dunque sono vincolate le (occorrenze delle) variabili che appaiano nel ‘raggio d’azione’ di un quantificatore (si intende però che ogni quantificatore può vincolare solo le occorrenze della variabile che lo segue immediatamente: in  $\forall x(x = y)$ , il quantificatore vincola la  $x$ , non la  $y$ ). In una qualsiasi formula, le (occorrenze delle) variabili che non sono vincolate si dicono *libere*. Ad esempio, sono vincolate le occorrenze di  $x$  in “ $(\forall x(x + 1 > x)) \wedge (\exists x(x > y))$ ”, mentre nella stessa formula è libera l’occorrenza di  $y$ . Attenzione: è possibile che nella stessa formula la stessa variabile abbia sia occorrenze libere che occorrenze vincolate: ad esempio, in “ $(\forall x(x + 1 > x)) \vee (x = 0)$ ” l’ultima occorrenza di  $x$  è libera, le altre sono vincolate (si noti però che tutte le occorrenze di  $x$  in “ $\forall x((x + 1 > x) \vee (x = 0))$ ” sono vincolate; attenzione alle parentesi!).

<sup>17</sup>l’origine del simbolo  $\exists$  è ovvia: richiama la ‘E’ iniziale della parola ‘Esiste’ o delle sue varianti in diverse lingue. Il simbolo  $\forall$  richiama invece la ‘A’ iniziale dell’inglese ‘All’ (o meglio, del tedesco ‘Alle’) che sta per ‘tutti’.

<sup>18</sup>cioè: introdotta da un quantificatore universale.

<sup>19</sup>non si tratta effettivamente di una congiunzione, perché si possono coniugare, usando il connettivo  $\wedge$ , solo un numero finito di formule per volta, invece le formule  $\varphi(a)$  sono, in generale, in numero infinito.

<sup>20</sup>questa frase vuol dire:  $\exists!$  è definito dal fatto che scelti comunque la formula  $\varphi$  e le variabili (distinte)  $x$  e  $y$  in modo che  $y$  non appaia in  $x$  la formula che segue è *valida*, nel senso [indicato sopra](#).

Benché le nozioni di occorrenze libere e vincolate siano qui state presentate solo per grandi linee, è bene farci attenzione per almeno due motivi. Il primo è che la (già promessa) definizione di formula chiusa dipende proprio dalle nozioni appena introdotte: *una formula è chiusa se e solo se non contiene variabili con occorrenze libere*. Vediamo così che formule come “ $\forall x(x > 1)$ ” e “ $\forall x((x+1 > x) \vee (x = 0))$ ”, che abbiamo incontrato poco sopra, sono proposizioni: in entrambe l'unica variabile che appare,  $x$ , ha solo occorrenze vincolate. Invece, la formula “ $\exists x(x > y)$ ” non è una proposizione, a causa dell'occorrenza libera di  $y$ . Possiamo modificare questa formula perché diventi chiusa? Certamente, il modo ovvio di farlo è quello di premettere un quantificatore che vincoli la variabile  $y$ : “ $\exists y(\exists x(x > y))$ ” e “ $\forall y(\exists x(x > y))$ ” sono proposizioni. Cosa significa tutto ciò in termini più semplici? Che se una formula contiene variabili che non sono state introdotte da quantificatori (in tutte le loro occorrenze) allora questa formula non ha un valore di verità e non ha senso stare a discutere sul fatto che sia vera o falsa. Per poterle attribuire un valore di verità dobbiamo prima ‘quantificare’, come si dice, le variabili libere che vi appaiono. Ovviamente abbiamo più modi di farlo, che portano, in genere, a formule molto diverse tra loro, come nel caso delle due formule  $\exists y(\exists x(x > y))$  e  $\forall y(\exists x(x > y))$  appena viste.

È bene insistere su questo punto: i quantificatori sono essenziali per la corretta espressione ed interpretazione delle formule in logica e più generalmente in matematica, ed è quindi importantissimo che siano sempre espressi, ed in modo non ambiguo.<sup>21</sup> Si invita chi legge a prendere la sana abitudine di farlo; la pratica di omettere o sottintendere quantificatori è una delle più frequenti cause di errori di ragionamento e, cosa forse più grave, di completa incomprensione di concetti matematici.

La seconda ragione per cui è bene cercare di familiarizzarsi con la nozione di occorrenza libera o vincolata di una variabile è che, nell'interpretazione di una formula, variabili libere e variabili vincolate giocano ruoli molto diversi. La questione è molto delicata e non facile da afferrare, non mi stupirei se il contenuto dei paragrafi seguenti risultasse poco comprensibile a chi legge, ma invito per lo meno a provare a rifletteci sopra.

Guardiamo alle sostituzioni. Non abbiamo potuto dare più che una idea intuitiva di come si effettui una sostituzione di un termine ad una variabile in una formula, ma una cosa che possiamo dire è che nelle sostituzioni si opera solo sulle occorrenze libere delle variabili. Se ci fermiamo a pensarci un attimo, questo non è strano. Infatti a nessuno (spero) verrebbe mai in mente di sostituire, ad esempio, 0 alla variabile  $x$  in “ $\forall x(x > y)$ ” ottenendo “ $\forall 0(x > y)$ ”, o magari “ $\forall x(0 > y)$ ”. Qualche esempio: nel linguaggio dell'aritmetica, consideriamo le tre formule  $\varphi$ ,  $\bar{\varphi}$  e  $\psi$ :

$$\varphi: "x > 1"; \quad \bar{\varphi}: "\forall x(x > 1)"; \quad \psi: "\forall x(x > 1) \wedge x = 7".$$

Sostituendo ad  $x$  il numero 3 in ciascuna delle formule otteniamo le formule

$$\varphi(3): "3 > 1"; \quad \bar{\varphi}(3) = \bar{\varphi}(x): "\forall x(x > 1)"; \quad \psi(3): "\forall x(x > 1) \wedge 3 = 7".$$

Cosa è successo? In ciascuna delle formule abbiamo sostituito 3 alle occorrenze libere di  $x$ , ma non a quelle vincolate; è questa la regola generale. Nelle formule chiuse non appaiono variabili libere, quindi non c'è nulla da sostituire e per questo le sostituzioni lasciano invariate le formule chiuse. Come si vede, questo è il caso che si è verificato per la formula (chiusa)  $\bar{\varphi}(x)$ . In  $\psi(x)$  l'unica occorrenza libera di  $x$  è l'ultima, quella che appare in “ $x = 7$ ”, quindi solo questa è stata sostituita.

Per dirla in modo grossolano ma comprensibile, una variabile libera è qualcosa che, nella formula, “rappresenta” un oggetto (e quindi può essere sostituita da un oggetto), una variabile vincolata invece no. Non per niente, per indicare una variabile con occorrenza vincolata si usa anche l'espressione ‘variabile muta’.

Infine, un po' di terminologia. Un *predicato unario* nella variabile  $x$  è formula che non contenga occorrenze libere di variabili diverse di  $x$ . Quindi, se  $\varphi$  è una formula e  $x$  è una variabile,  $\forall x(\varphi)$  è una proposizione esattamente quando  $\varphi$  è un predicato unario in  $x$ ; lo stesso vale per  $\exists x(\varphi)$  e  $\exists!x(\varphi)$ . Similmente, si dice che la formula  $\varphi$  è un *predicato binario* quando in essa appaiono al più due variabili con occorrenze libere<sup>22</sup>, un *predicato ternario* è una formula in cui appaiono al più tre variabili con occorrenze libere, e così via.

<sup>21</sup>ad esempio, andrebbero evitate espressioni del tipo: “ $f(x) < 4$ , con  $x > 0$ ” dove chi legge deve, per bene che vada, tirare ad indovinare se quel “con” rappresenta un quantificatore esistenziale o uno universale. Analogamente, in “..., per  $x > 0$ ”, non sempre è chiaro se quel “per” vada inteso come “per ogni” o “per almeno un”. Il senso logico delle frasi, come si vede, cambia se si cambia l'interpretazione.

<sup>22</sup>attenzione: non due *occorrenze*, ma due variabili, con un numero arbitrario di occorrenze. Ad esempio, la formula “ $(x < y) \vee ((x = y) \Rightarrow (x > 7))$ ”, in cui  $x$  appare tre volte, è un predicato binario in  $x$  e  $y$  (nel linguaggio dell'aritmetica).

**Osservazioni.**

**G.1.** Torniamo sulla equivalenza che abbiamo dato come **descrizione del quantificatore  $\exists!$** . Siano  $\varphi$  una formula e  $x, y$  due variabili, e assumiamo che  $y$  non appaia in  $\varphi$ . Se chiamiamo  $\psi(x, y)$  la formula  $\varphi(y) \Leftrightarrow y = x$ , possiamo riscrivere l'equivalenza come  $\exists!x(\varphi(x)) \Leftrightarrow \exists x(\forall y(\psi(x, y)))$ . Vogliamo esaminare il membro a destra di questa equivalenza. Per semplificare il discorso, supponiamo che  $\varphi$  sia un predicato unario in  $x$ , quindi che  $\exists x(\forall y(\psi(x, y)))$  sia una proposizione; il ragionamento è analogo nel caso generale. Quando è che questa proposizione è vera? Esattamente quando esiste *almeno* un  $a$  per il quale sia vera la formula (che è anch'essa chiusa)  $\forall y(\psi(a, y))$ ; come sappiamo questo equivale a dire che è vera  $\psi(a, b)$ , cioè la formula  $\varphi(b) \Leftrightarrow b = a$ , per ogni possibile scelta di  $b$ . Tra le possibili scelte per  $b$  c'è anche  $a$ ; la formula diventa in questo caso particolare  $\varphi(a) \Leftrightarrow a = a$ . Poiché  $a = a$  è vera, questa equivale a  $\varphi(a)$ . Se invece scegliamo come  $b$  un qualsiasi oggetto diverso da  $a$ , allora  $b = a$  è falsa, quindi  $\varphi(b) \Leftrightarrow b = a$  equivale alla negazione di  $\varphi(b)$ .

In definitiva, abbiamo mostrato che la formula  $\exists x(\forall y(\psi(x, y)))$  è vera se e solo se esiste un  $a$  per il quale è vera  $\varphi(a)$  e, contemporaneamente, è falsa  $\varphi(b)$  per ogni  $b$  diverso da  $a$ . Questo è precisamente quello che volevamo esprimere col quantificatore  $\exists!$ ; è quindi giustificata l'idea di descrivere formalmente questo quantificatore come abbiamo fatto.

**G.2.** Non abbiamo descritto in dettaglio le sostituzioni, abbiamo solo avvertito che la nozione è solo ingannevolmente semplice. Un esempio può dare l'idea delle difficoltà che possono sorgere (ma, niente paura, si risolvono). È lecito sostituire a variabili altre variabili. Consideriamo la formula  $\exists x(x \neq y)$  nella variabili  $x$  (vincolata) e  $y$  (libera). Cosa succederebbe se sostituissimo 'meccanicamente'  $y$  con  $x$  in questa formula?

**G.3.** Questa osservazione fornisce un suggerimento pratico, mirato a semplificare la scrittura delle formule. Una regola (molto intuitiva) del calcolo dei predici afferma che se si "cambia nome" alle variabili vincolate in una formula si ottiene una formula equivalente. Più precisamente, data una formula  $\varphi$  in cui appare una variabile  $x$ , se  $y$  è una variabile che non appare in  $\varphi$ , la formula ottenuta scrivendo  $y$  al posto di  $x$  in ogni occorrenza vincolata di  $x$  in  $\varphi$  è equivalente a  $\varphi$ . Per esempio, " $\forall x(x + 1 > x)$ " e " $\forall y(y + 1 > y)$ " sono equivalenti. Usando questa regola, è possibile riscrivere in modo equivalente qualsiasi formula in modo da evitare che la stessa variabile appaia sia libera che vincolata, con gran vantaggio per la chiarezza. Ad esempio, la già menzionata formula " $(\forall x(x + 1 > x)) \vee (x = 0)$ " dell'aritmetica si potrebbe equivalentemente riscrivere come " $(\forall y(y + 1 > y)) \vee (x = 0)$ ", sicuramente più facile da leggere.

**Quantificatori ristretti.** Nella pratica matematica si incontrano con gran frequenza espressioni del tipo " $(\forall x \in S)(\varphi)$ " o " $(\exists x > 0)(\varphi)$ " (le parentesi non sono tutte necessarie, ma rendono le formule più facili da leggere; come si sarà immaginato qui  $\varphi$  è una formula e  $x$  una variabile, e  $S$  è un insieme), in cui il quantificatore è accompagnato da una condizione che limita l'ambiente in cui la variabile possa assumere i suoi valori. Queste espressioni hanno ovvie interpretazioni, ma è bene sapere che sono semplicemente abbreviazioni di formule in cui i quantificatori sono usati nel modo indicato nella sezione precedente, ed è bene sapere di quali formule sono abbreviazioni. La prima formula può essere definita in questo modo:

$$(\forall x \in S)(\varphi) : \Leftrightarrow \forall x(x \in S \Rightarrow \varphi)$$

(i due punti che precedono  $\Leftrightarrow$  ci ricordano solo che questa equivalenza, o meglio l'affermazione che vale questa equivalenza, è stabilita come definizione dell'espressione a sinistra). Come spesso accade, non essersi accortati di una idea intuitiva ma aver cercato una definizione precisa non è un atto di pignoleria fine a se stesso, ma comporta un utile vantaggio. In questo caso, ci permette di chiarire in modo molto semplice un punto che spesso sfugge agli studenti: cosa accade quando  $S$  è l'insieme vuoto? La risposta è:

per ogni predicato unario  $\varphi$  nella variabile  $x$ , la proposizione  $(\forall x \in \emptyset)(\varphi)$  è vera.<sup>23</sup>

Come mai? Stando alla nostra definizione " $(\forall x \in \emptyset)(\varphi)$ " significa " $\forall x(x \in \emptyset \Rightarrow \varphi)$ ". Ora, qualunque sia l'oggetto  $a$ , la formula  $a \in \emptyset$  è falsa (l'insieme vuoto non ha elementi: è questa la sua definizione) quindi l'implicazione " $a \in \emptyset \Rightarrow \varphi(a)$ " ha l'antecedente falso e quindi è vera. Dunque, se sostituiamo  $a$  ad  $x$  in " $x \in \emptyset \Rightarrow \varphi$ " otteniamo certamente una formula vera. Pertanto " $\forall x(x \in \emptyset \Rightarrow \varphi)$ ", ovvero

---

<sup>23</sup>ricordiamo che se la formula  $\varphi$  non fosse un predicato unario in  $x$ , cioè se  $\varphi$  contenesse una variabile libera diversa da  $x$ , allora  $(\forall x \in \emptyset)(\varphi)$  non sarebbe una proposizione, quindi non sarebbe né vera né falsa.

$(\forall x \in \emptyset)(\varphi)$ , è vera, come si voleva dimostrare. Possiamo a questo punto dire che, a meno che non esistano cavalli<sup>24</sup> verdi, la frase “ogni cavallo verde ha otto zampe” è vera.

Discorso analogo vale per altre restrizioni che possono essere imposte alla variabile quantificata. Ad esempio,  $(\forall x > 0)(\varphi)$  significa  $\forall x(x > 0 \Rightarrow \varphi)$ .

Se cambiamo quantificatore la definizione è diversa:

$$(\exists x \in S)(\varphi) : \iff \exists x((x \in S) \wedge \varphi)$$

e qui non ci dovrebbero essere difficoltà: “esiste  $x$  in  $S$  tale che ...” significa proprio “esiste  $x$  tale che  $x$  sia in  $S$  e ...”. Ovviamente, nella solita ipotesi che  $\varphi$  sia un predicato unario in  $x$ , questa formula è sicuramente una proposizione falsa quando  $S = \emptyset$ .

Abbiamo introdotto i quantificatori  $\forall$  ed  $\exists$  suggerendo un’analoga tra essi ed i connettivi  $\wedge$  e  $\vee$ , cioè che i quantificatori in qualche modo corrispondano a forme più generali di congiunzione (nel caso del quantificatore universale) e disgiunzione (per quello esistenziale). Questa analogia si può effettivamente rendere precisa e verificare nel caso dei quantificatori ristretti ad insiemi finiti e non vuoti. Se  $S$  è appunto un insieme finito e  $S \neq \emptyset$ , se  $\varphi$  è una formula e  $x$  una variabile, è chiaro che la formula  $(\forall x \in S)(\varphi(x))$  è equivalente a  $\bigwedge_{a \in S} \varphi(a)$ , cioè a  $\varphi(a_1) \wedge \varphi(a_2) \wedge \dots \wedge \varphi(a_k)$ , dove  $a_1, \dots, a_k$  sono gli elementi di  $S$ , mentre  $(\exists x \in S)(\varphi(x))$  è equivalente a  $\bigvee_{a \in S} \varphi(a)$ .<sup>25</sup>

Si può, a questo punto, tornare all’esempio della frase “per ogni numero intero  $x$  compreso tra 1 e 3 si ha che se  $x > 2$  allora  $x > 1$ ” discussa nella sezione in cui è stato introdotto il connettivo di implicazione. Come si può ora riconoscere, questa frase non è altro che un modo per rendere verbalmente la formula  $(\forall x \in \{1, 2, 3\})(x > 2 \Rightarrow x > 1)$ ; questa è una formula introdotta da un quantificatore universale ristretto ad un insieme di tre elementi, quindi si riduce ad una congiunzione tra tre formule, e sotto questo aspetto l’avevamo studiata.

## 6. QUALCHE REGOLA D’USO

Esiste un gran numero di regole del calcolo dei predicati che permettono di manipolare formule contenenti quantificatori. Si tratta per lo più di regole estremamente intuitive; una è quella data nell’Osservazione G.3, ne vedremo altre. Spesso queste regole sono enunciate dichiarando la validità di determinate implicazioni. Anche in questa sezione, ma non lo ripeteremo ogni volta, le lettere  $x$ ,  $y$  e  $z$  indicano sempre variabili,  $\varphi$  e  $\psi$  sono invece formule.

**Quantificatori multipli.** Innanzitutto, può capitare di avere più quantificatori consecutivi; un esempio lo abbiamo già visto con la formula a secondo membro della equivalenza che descrive formalmente  $\exists!$ . Abbiamo formule del tipo  $\forall x(\forall y(\dots(\varphi)\dots))$  o  $\exists x(\exists y(\dots(\varphi)\dots))$ , in cui è lo stesso quantificatore a ripetersi; in questi casi l’ordine in cui appaiono i quantificatori è irrilevante nel senso che, ad esempio,  $\forall x(\forall y(\varphi))$  e  $\forall y(\forall x(\varphi))$  sono equivalenti. Si usa scrivere, per brevità,  $\forall x, y, \dots, z (\varphi)$  invece di  $\forall x(\forall y(\dots \forall z(\varphi)\dots))$  e  $\exists x, y, \dots, z (\varphi)$  invece di  $\exists x(\exists y(\dots \exists z(\varphi)\dots))$ . Diverso è il caso in cui appaiono sia il quantificatore esistenziale che quello universale. Le formule “ $\forall x(\exists y(\varphi))$ ” e “ $\exists y(\forall x(\varphi))$ ” non sono in generale equivalenti. La prima afferma che, scelto comunque un termine  $a$ , ne esiste almeno uno,  $b$ , dipendente, in generale, dalla scelta di  $a$ , per il quale si abbia  $\varphi(a, b)$ . La seconda formula dice qualcosa in più: che si ha la stessa situazione ma, questa volta, si può scegliere  $b$  indipendentemente dalla scelta di  $a$ : esiste un particolare  $b$  per il quale si abbia  $\varphi(a, b)$  per ogni possibile scelta di  $a$ . Dunque, vale sempre l’implicazione

$$\exists y(\forall x(\varphi)) \implies \forall x(\exists y(\varphi))$$

ma, in generale, non vale l’implicazione inversa. Un esempio può aiutare: nel linguaggio dell’aritmetica, sia  $\varphi(x, y)$  la formula  $x < y$ . La prima delle nostre formule diventa

$$\forall x(\exists y(x < y)),$$

---

<sup>24</sup>si intende: cavalli *vivi*.

<sup>25</sup>per una convenzione di uso universale, se  $S$  ha un solo elemento, cioè  $k = 1$  e  $S = \{a_1\}$ , sia  $\bigwedge_{a \in S} \varphi(a)$  che  $\bigvee_{a \in S} \varphi(a)$  valgono  $\varphi(a_1)$ . In verità la stessa convenzione permette di estendere questa notazione anche al caso in cui  $S = \emptyset$ , stabilendo che  $\bigwedge_{a \in \emptyset} \varphi(a)$  e  $\bigvee_{a \in \emptyset} \varphi(a)$  indicano una formula vera ed una falsa, rispettivamente. Con questa ulteriore convenzione le equivalenze  $((\forall x \in S)(\varphi(x))) \Leftrightarrow \bigwedge_{a \in S} \varphi(a)$  e  $((\exists x \in S)(\varphi(x))) \Leftrightarrow \bigvee_{a \in S} \varphi(a)$  continuano a valere, anche nel caso in cui  $S = \emptyset$ .

che afferma che per ogni numero esiste un numero più grande. Questa è una proposizione vera: se  $a$  è un numero intero,  $a + 1$  è un numero intero maggiore di  $a$ , quindi  $\varphi(a, a + 1)$  è vera.<sup>26</sup> La seconda formula è invece

$$\exists y(\forall x(x < y)),$$

che afferma che esiste un intero (quello che andrebbe sostituito ad  $y$ ) maggiore di tutti gli interi; questa è una proposizione falsa.

**Negazione di quantificatori.** Come si nega una formula universale? E come si nega una formula esistenziale? Dovrebbe bastare il buon senso a suggerirlo: per stabilire che sia falsa la frase “ogni cittadino italiano si chiama Mario” basta osservare che esiste qualche cittadino italiano che non si chiama Mario; anche se ce ne sono alcuni che effettivamente si chiamano Mario la frase è ugualmente falsa, perché *non tutti* hanno quel nome. Sarebbe un errore pensare che per rendere falsa la frase in questione bisognerebbe che *nessuno* dei cittadini italiani si chiamasse Mario. Questo esempio suggerisce che la negazione di una frase universale sia una frase esistenziale (dove è negata la formula oggetto della quantificazione). È proprio così; per ogni formula  $\varphi$  ed ogni variabile  $x$  vale:

$$(\neg(\forall x(\varphi))) \iff (\exists x(\neg\varphi)). \quad (\text{negazione di formule universali})$$

Simmetricamente, pensiamo che la frase “esiste un cittadino italiano di nome Xwas” sia falsa, non perché esiste un cittadino italiano che non si chiama Xwas, ma perché *nessun* cittadino italiano si chiama Xwas, o, per dirla in modo più utile ai nostri scopi, anche se meno naturale, *ogni* cittadino italiano *non* si chiama Xwas.<sup>28</sup> La regola di negazione per le formule esistenziali è, infatti:

$$(\neg(\exists x(\varphi))) \iff (\forall x(\neg\varphi)). \quad (\text{negazione di formule esistenziali})$$

Vale la pena di osservare che questa regola segue dalla precedente e dalle tautologie della **negazione** e della **commutatività** dell’equivalenza. Infatti, “ $(\neg(\exists x(\varphi))) \iff (\forall x(\neg\varphi))$ ” equivale, per queste tautologie, a “ $(\exists x(\varphi)) \iff (\neg(\forall x(\neg\varphi)))$ ” e quindi a “ $(\neg(\forall x(\neg\varphi))) \iff (\exists x(\varphi))$ ”. Questa (per la tautologia della **doppia negazione**) non è altro che la regola per la negazione delle formule universali applicata con  $\neg\varphi$  al posto di  $\varphi$ .

Se torniamo all’analogia tra i quantificatori  $\forall$  ed  $\exists$  ed i connettivi  $\wedge$  e  $\vee$ , possiamo pensare a queste regole di negazione come all’analogo delle **leggi di De Morgan**.

Notiamo che queste regole di negazione stabiliscono anche l’interdipendenza di  $\forall$  ed  $\exists$ , nel senso che mostrano come l’uno dei due si possa definire in termini dell’altro; ad esempio, potremmo assumere dato  $\forall$  e definire  $\exists$  usando l’equivalenza  $(\exists x(\varphi)) \Leftrightarrow (\neg(\forall x(\neg\varphi)))$ . La situazione, come si vede, è simile a quella dei connettivi proposizionali: abbiamo introdotto due simboli ( $\forall$  e  $\exists$ ) ma ci siamo accorti che, volendo, potremmo fare a meno di uno dei due.

Non sorprendentemente, per formule con quantificatori ristretti abbiamo regole di negazione simili a quelle per i quantificatori non ristretti (le notazioni sono quelle solite; in particolare,  $S$  indica un insieme):

$$\neg(\forall x \in S)(\varphi) \iff \exists(x \in S)(\neg\varphi) \quad \text{e} \quad \neg(\exists x \in S)(\varphi) \iff \forall(x \in S)(\neg\varphi).$$

Il senso è chiaro, ma è utile e istruttivo verificare queste formule. Per la prima:  $(\forall x \in S)(\varphi)$  significa  $\forall x(x \in S \Rightarrow \varphi)$ , la cui negazione è  $\exists x(\neg(x \in S \Rightarrow \varphi))$ . Ricordiamo **come si nega un’implicazione**: affermando l’antecedente e contemporaneamente negando il conseguente. Quindi  $\neg(\forall x \in S)(\varphi)$  equivale a  $\exists x((x \in S) \wedge (\neg\varphi))$ . Ma questa formula, come abbiamo visto sopra, è proprio quella che viene abbreviata con  $\exists(x \in S)(\neg\varphi)$ . La verifica è così completa. La seconda formula si può dimostrare dalla prima (analogamente a quanto fatto nel caso dei quantificatori non ristretti) oppure in modo diretto, come si suggerisce di fare in [uno dei prossimi esercizi](#).

<sup>26</sup> Seguiamo qui una convenzione standard: avendo indicato la formula  $\varphi$  come  $\varphi(x, y)$ , abbiamo specificato l’ordine in cui consideriamo le variabili. Dunque  $\varphi(a, a + 1)$  sarà la formula ottenuta sostituendo  $a$  ad  $x$  e  $a + 1$  ad  $y$ , non viceversa.

<sup>27</sup> Stiamo facendo largo uso di parentesi, sperando che questo aiuti nella lettura. Potremmo però anche farne a meno. Ad esempio, questa formula si potrebbe anche scrivere  $\neg\forall x(\varphi) \iff \exists x(\neg\varphi)$ , senza nessuna ambiguità; si vedano le tautologie della **negazione dell’equivalenza** per il ruolo di  $\neg$  nella sua prima occorrenza.

<sup>28</sup> Spero di non essere smentito, su questo punto, da un’indagine anagrafica.

**Un errore da evitare.** Qualunque sia il termine  $a$ , due (ovvie) regole<sup>29</sup> stabiliscono la catena di implicazioni:

$$(\forall x(\varphi(x))) \Rightarrow \varphi(a) \Rightarrow (\exists x(\varphi(x))),$$

da cui segue  $(\forall x(\varphi(x))) \Rightarrow (\exists x(\varphi(x)))$  (a condizione che si ammetta, come in genere si fa, che esista almeno un oggetto a cui il linguaggio si riferisce). Nel caso in cui i quantificatori siano ristretti la situazione può essere diversa. L'implicazione  $((\forall x \in S)(\varphi(x))) \Rightarrow ((\exists x \in S)(\varphi(x)))$  vale certamente se  $S$  è un insieme non vuoto, ma non se  $S$  è l'insieme vuoto. Per convincercene, consideriamo il caso in cui  $\varphi$  sia un predicato unario in  $x$ . Se  $S = \emptyset$  l'antecedente  $(\forall x \in \emptyset)(\varphi(x))$  della nostra implicazione è vero (si veda la discussione su queste formule nella sezione sui quantificatori ristretti) ed il conseguente  $(\exists x \in \emptyset)(\varphi(x))$  è falso, quindi l'implicazione è falsa.

Capita non tanto di rado di trovare errori di ragionamento dovuti proprio a questa disattenzione: dal fatto che tutti gli elementi di un insieme abbiano una certa proprietà si deduce l'esistenza di almeno un elemento con quella proprietà. Questo passaggio non è lecito a meno di non essersi assicurati che l'insieme in questione non è vuoto. Per esempio, abbiamo detto che ogni cavallo verde ha otto zampe, da questo non possiamo certamente dedurre che esistano cavalli verdi con otto zampe!

### Esercizi.

**H.1.** Vero o falso? E perché? Questo è un esercizio di corretta lettura ed interpretazione di formule.

- (a)  $(\forall x \in \mathbb{N})(x + 1 < x \Rightarrow x^2 = 1)$ .
- (b)  $\exists x \in \mathbb{N}(\forall y \in \mathbb{N}(x \leq y))$ .
- (c)  $\forall x \in \mathbb{N}(\exists y \in \mathbb{N}(x < y))$ .
- (d)  $\forall x \in \mathbb{N}(\exists y \in \mathbb{N}((x = y + 1) \Rightarrow (x < y)))$ .
- (e)  $\exists x \in \mathbb{N}(\forall y \in \mathbb{N}((x < y) \vee (y < x) \vee (y = 11)))$ .
- (f)  $\exists x \in \mathbb{N}(\forall y \in \mathbb{Z}((x \neq y) \Rightarrow (x < y)))$ .
- (g) Ogni numero reale il cui quadrato sia negativo è maggiore di  $10^{327}$ .

**H.2.** Una regola (molto intuitiva) del calcolo dei prediciati stabilisce l'equivalenza

$$(\forall x(\varphi \wedge \psi)) \iff ((\forall x(\varphi)) \wedge (\forall x(\psi))),$$

qualsiasi siano le formule  $\varphi$  e  $\psi$  (ovviamente  $x$  indica una variabile). Oppure mi sbaglio? Anche questo è un esercizio di lettura! Può essere utile pensare a frasi come “ogni giorno mangio una pizza e vado al cinema” e “ogni giorno mangio una pizza e ogni giorno vado al cinema”. Con le stesse notazioni, confrontare tra loro le formule “ $\forall x(\varphi \vee \psi)$ ” e “ $(\forall x(\varphi)) \vee (\forall x(\psi))$ ”.

Ripetere l'esercizio sostituendo, in tutte le formule, “ $\forall$ ” con “ $\exists$ ”.

**H.3.** Verificare (in modo diretto) la formula  $\neg(\exists x \in S)(\varphi) \iff \forall(x \in S)(\neg\varphi)$ .

**H.4.** Si scriva la negazione di  $\exists!x(\varphi)$ . Sono possibili più risposte, diverse nella forma ma equivalenti tra loro.

**H.5.** Si neghi ciascuna delle le formule (le notazioni sono le solite):

- (a)  $\forall x(\exists y(\varphi(x, y) \Rightarrow \psi(x, y)))$ ;
- (b)  $\exists x(\varphi(x) \wedge \forall y(\neg\psi(x, y)))$ ;
- (c)  $\forall x, y(\exists z(z \neq y \wedge \varphi(x, z)))$ ;

**H.6.** Si neghi ciascuna delle frasi:

- (a) Ogni volta che vedo Astolfo, litighiamo.
- (b) Una volta ho visto Astolfo ed ho bevuto un caffé.
- (c) Tutti i giorni della prossima settimana andrò al cinema, ed uno di quei giorni andrò in pizzeria.

## 7. INSIEMI

Sia  $\varphi = \varphi(x)$  un predicato unario nella variabile  $x$ . Si indica col simbolo  $\{x \mid \varphi\}$  la totalità (potremmo anche dire: la collezione, la classe; stiamo usando questi termini in modo del tutto informale) degli oggetti  $a$  che, sostituiti ad  $x$  in  $\varphi$ , rendono  $\varphi$  vera (vale a dire: tali che  $\varphi(a)$  sia una formula vera). Questa totalità si chiama anche l'estensione di  $\varphi$ . Si può pensare a  $\varphi$  come espressione di una ‘proprietà’ che un oggetto può soddisfare oppure no, allora la sua estensione  $\{x \mid \varphi\}$  è la totalità degli oggetti che soddisfano (verificano, hanno) la proprietà espressa da  $\varphi$ . Ad esempio, la formula “ $(x \in \mathbb{N}) \wedge (x < 3)$ ”,

<sup>29</sup>la prima delle quali si chiama *regola di specializzazione*

esprime la proprietà di “essere un numero naturale minore di 3”, e  $\{x \mid (x \in \mathbb{N}) \wedge (x < 3)\}$  è costituito dai numeri 0, 1 e 2.

La teoria degli insiemi tratta di enti matematici, appunto gli insiemi, che formalizzano l’idea intuitiva di ‘aggregato di oggetti’; gli assiomi di questa teoria regolano in modo preciso il modo in cui su questi enti si può operare. Un tentativo (quello del logico tedesco Gottlob Frege) di fondare l’intera teoria sull’idea che l’estensione di ogni predicato unario si possa considerare come insieme fallì molto presto, non appena si scoprì che questa assunzione porta necessariamente a contraddizioni.<sup>30</sup> In altri termini: non sempre l’estensione  $\{x \mid \varphi(x)\}$  di un predicato  $\varphi$  (unario, in  $x$ ) è un insieme. Detto in modo ancora diverso, esistono ‘proprietà’ perfettamente ragionevoli e ben definite (cioè espresse da un predicato unario) tali che, sfortunatamente, non esiste l’*insieme* degli oggetti che le verificano. Ad esempio, tutti gli studenti che sono stati sottoposti ad una qualche infarinatura di teoria degli insiemi (come questa!) dovrebbero sapere che *non esiste l’insieme di tutti gli insiemi*; si vedano a questo proposito l’[Esempio I.2](#) e l’[Esercizio I.3](#).

Qualcosa dell’idea di Frege, però, si salva. Dato un predicato unario  $\varphi$ , se proviamo a selezionare tutti gli oggetti che verificano  $\varphi$  è possibile, lo abbiamo appena detto, che non si ottenga un insieme; ma se noi abbiamo già a disposizione un insieme  $S$  e limitiamo la selezione ai soli elementi di  $S$  (lasciando perdere tutto ciò che non appartenga ad  $S$ ), allora sicuramente otteniamo un insieme: l’insieme degli elementi di  $S$  che verificano  $\varphi$ . Ce lo assicura uno degli assiomi della teoria degli insiemi, l’*assioma di separazione* (o di *comprendizione*; più precisamente, si tratta di uno schema di assiomi, ma non entriamo in questa sottigliezza). In modo più esplicito: dati un insieme  $S$  ed un predicato unario  $\varphi$  nella variabile  $x$ , la formula “ $(x \in S) \wedge \varphi$ ” è ancora un predicato unario in  $x$ ;<sup>31</sup> l’assioma di separazione dice che, in queste circostanze, l’estensione  $\{x \mid (x \in S) \wedge \varphi(x)\}$  di questo predicato è un insieme. Per indicare questo insieme si usa, in genere, una notazione più compatta:  $\{x \in S \mid \varphi(x)\}$ . L’insieme così ottenuto è ovviamente una parte di  $S$ .

Un altro assioma che conviene menzionare è l’assioma di *estensionalità*. Questo assioma stabilisce che gli insiemi sono completamente determinati dai loro elementi, ovvero: dati un insieme  $A$  ed un insieme  $B$ , si ha  $A = B$  se e solo se  $A$  e  $B$  hanno esattamente gli stessi elementi—si veda l’[Esercizio I.5](#). Abbiamo implicitamente fatto uso di questo assioma quando abbiamo descritto  $\{x \mid \varphi(x)\}$  (quando è un insieme) e  $\{x \in S \mid \varphi(x)\}$  specificando solo quali sono i loro elementi.

Torniamo sul significato di queste espressioni. Trattiamo come oggetti matematici ‘veri e propri’ gli insiemi, ma non le estensioni di predicati che non siano insiemi. Quindi, continuando ad usare le notazioni che abbiamo introdotto nei paragrafi precedenti, per noi  $\{x \mid \varphi(x)\}$  esiste come oggetto della matematica se è un insieme, non esiste altrimenti. In effetti, più in generale, nella teoria degli insiemi standard si assume abitualmente che *non esistano enti matematici che non siano insiemi*.<sup>32</sup> Scriviamo dunque formule come  $A = \{x \mid \varphi(x)\}$  solo nel caso in cui  $\{x \mid \varphi(x)\}$  sia un insieme. In questo caso,

$$\text{l’uguaglianza } A = \{x \mid \varphi(x)\} \quad \text{equivale a: } \forall x(x \in A \Leftrightarrow \varphi(x)).$$

Invece la formula  $A = \{x \in S \mid \varphi(x)\}$  ha sempre senso, perché il secondo membro è un insieme, e

$$\text{l’uguaglianza } A = \{x \in S \mid \varphi(x)\} \quad \text{equivale a: } \forall x(x \in A \Leftrightarrow ((x \in S) \wedge \varphi(x))).$$

### Esempi, Osservazioni, Esercizi (alcuni non facili).

**I.1.** Usando l’assioma di estensionalità verificare che esiste solo un insieme vuoto.

**I.2.** Abbiamo detto che se  $\varphi$  è un predicato unario, non necessariamente esiste l’insieme  $\{x \mid \varphi(x)\}$  degli oggetti che verificano  $\varphi$ . Per convincerci di questo fatto, esaminiamo un esempio. Scegliamo come  $\varphi$  la formula  $x \notin x$  e supponiamo che esista l’insieme  $\{x \mid x \notin x\}$ , che possiamo chiamare  $R$ . Stando al significato che abbiamo stabilito per questa notazione, abbiamo:  $\forall x(x \in R \Leftrightarrow x \notin x)$ . La [regola di specializzazione](#), applicata sostituendo di  $R$  a  $x$ , fornisce allora  $R \in R \Leftrightarrow R \notin R$ . Questa è evidentemente una contraddizione; dobbiamo concludere che l’insieme  $\{x \mid x \notin x\}$  non esiste.

Questo esempio contiene la cosiddetta *Antinomia* (o *Paradosso*) di Russell (da ciò l’uso della lettera  $R$ ) ed ha una bella ed istruttiva storia alle spalle (anzi, [almeno due](#)).

<sup>30</sup>Stiamo rendendo breve una storia che è molto più lunga e complessa. In particolare, il primo sistema di assiomi per la teoria degli insiemi, quello di Zermelo (1908), è storicamente successivo all’emersione delle contraddizioni a cui stiamo accennando.

<sup>31</sup>Come già capitato in casi analoghi, ci stiamo prendendo una piccola libertà: trattiamo in questa formula  $S$  come un simbolo di costante. Per una versione meglio formulata dell’assioma di separazione si veda l’[Esercizio I.4](#).

<sup>32</sup>Esistono teorie degli insiemi alternative, un po’ più sofisticate, nelle quali le cose non stanno così e si può dare un significato matematico preciso a  $\{x \mid \varphi(x)\}$  anche nel caso in cui questo non sia un insieme.

**I.3.** Usando l'osservazione precedente (I.2) e l'**assioma di separazione**, dimostrare che non esiste l'insieme di tutti gli insiemi. Cosa sappiamo dire su  $\{x \mid x = x\}$ ?

**I.4.** Sia  $\varphi$  un predicato unario in  $x$ . La formula  $\forall y \exists z \forall x ((x \in z) \Leftrightarrow ((x \in y) \wedge \varphi))$  è uno dei modi per esprimere l'**assioma di separazione** ‘applicato’ a  $\varphi$  (per meglio dire: l'istanza dell'assioma di separazione per  $\varphi$ ). Verificarlo; è un ulteriore esercizio di lettura.

**I.5.** Assumendo che non esistano oggetti che non siano insiemi l'**assioma di estensionalità** è invece espresso dalla formula:  $(\forall y, z)(y = z \Leftrightarrow (\forall x(x \in y \Leftrightarrow x \in z)))$ . Verificarlo.

**Formule insiemistiche.** La vaga ed imperfetta corrispondenza tra prediciati unari ed insiemi, di cui abbiamo parlato nella sezione precedente, può essere comunque usata per tradurre risultati del calcolo proposizionale (come la validità di tautologie) in formule della teoria degli insiemi. Questa sorta di traduzione si può ottenere facendo corrispondere relazioni o operazioni insiemistiche a connettivi proposizionali, come vedremo con diversi esempi.

Come punto di partenza, osserviamo che l'idea di estensione di un predicato, discussa nella sezione precedente, si può in un certo senso invertire. Infatti, ogni insieme  $A$  si può considerare come l'estensione del predicato “ $x \in A$ ”, vale a dire:  $A = \{x \mid x \in A\}$ .<sup>33</sup> Ora, se  $A$  e  $B$  sono insiemi, l'**assioma di estensionalità** ci dice che vale:

$$A = B \Leftrightarrow \forall x(x \in A \Leftrightarrow x \in B),$$

mentre, per definizione di inclusione,

$$A \subseteq B \Leftrightarrow \forall x(x \in A \Rightarrow x \in B).$$

Queste due formule suggeriscono che, nello stesso senso, informale in cui facciamo corrispondere (ovvero ‘traduciamo’) i prediciati “ $x \in A$ ” e “ $x \in B$ ” con  $A$  e  $B$ , possiamo far corrispondere i connettivi  $\Leftrightarrow$  e  $\Rightarrow$  ai simboli di uguaglianza e di inclusione tra insiemi. Ci vuole poco a convincersi di come, nella stessa ottica, la **tautologia della doppia implicazione** si traduca nella ben nota regolaletta insiemistica della doppia inclusione. A titolo di esempio, verifichiamolo in dettaglio, senza timore di essere troppo pignoli. La tautologia assicura che “ $x \in A \Leftrightarrow x \in B$ ” sia equivalente a “ $(x \in A \Rightarrow x \in B) \wedge (x \in B \Rightarrow x \in A)$ ”. Inoltre, una delle tante regole del calcolo dei prediciati (si veda l'**Esercizio H.2**) stabilisce l'equivalenza

$$(\forall x(\varphi \wedge \psi)) \Leftrightarrow ((\forall x(\varphi)) \wedge (\forall x(\psi))). \quad (*)$$

qualsiasi siano le formule  $\varphi$  e  $\psi$ . Abbiamo allora le equivalenze (alcune scritte in verticale, questa volta, e giustificate da un commento a destra):

$$\begin{aligned} A = B &\Leftrightarrow \forall x(x \in A \Leftrightarrow x \in B) \\ &\Downarrow && \text{(tautologia)} \\ \forall x((x \in A \Rightarrow x \in B) \wedge (x \in B \Rightarrow x \in A)) && \\ &\Downarrow && \text{(regola (*))} \\ (\forall x(x \in A \Rightarrow x \in B)) \wedge (\forall x(x \in B \Rightarrow x \in A)) && \\ &\Downarrow && \text{(definizione di inclusione)} \\ (A \subseteq B) \wedge (B \subseteq A). && \end{aligned}$$

Quindi, e non è una sorpresa,  $A = B$  se e solo se  $A \subseteq B$  e  $B \subseteq A$ ; questa è la regola della doppia inclusione. In modo analogo, la tautologia della **transitività dell'implicazione** fornisce la *transitività dell'inclusione*:

$$(\forall A, B, C)((A \subseteq B) \wedge (B \subseteq C)) \Rightarrow (A \subseteq C);$$

è un utile esercizio verificarlo in dettaglio.

Abbiamo fatto corrispondere i simboli di uguaglianza e inclusione ai connettivi di equivalenza e di implicazione; è chiaro cosa si possa far corrispondere ai connettivi di congiunzione e disgiunzione: le operazioni di intersezione e di unione (binarie). Questo perché, scelti comunque gli insiemi  $A$  e  $B$ , si ha

$$\forall x(x \in A \cap B \Leftrightarrow ((x \in A) \wedge (x \in B))) \quad \text{e} \quad \forall x(x \in A \cup B \Leftrightarrow ((x \in A) \vee (x \in B))),$$

<sup>33</sup>ovviamente questa non sarebbe accettata come una vera e propria descrizione di  $A$ . Il discorso su questo punto andrebbe approfondito, ma così si andrebbe molto al là dei nostri scopi. Osserviamo ancora (vedi la nota 31) che nella formula “ $x \in A$ ” trattiamo  $A$  come una costante, similmente faremo in formule analoghe, dove altri insiemi appaiono al posto di  $A$ .

quindi il predicato di appartenenza ad  $A \cap B$  equivale alla congiunzione del predicato di appartenenza ad  $A$  e di quello di appartenenza ad  $B$ , mentre il predicato di appartenenza ad  $A \cup B$  equivale alla disgiunzione di questi due.<sup>34</sup> Dalle tautologie di **idempotenza**, **commutatività** e **associatività** (pag. 7) e dalle **leggi distributive** (pag. 8) per  $\wedge$  e  $\vee$  si ottengono dunque le analoghe proprietà per le operazioni di unione e intersezione tra insiemi: per ogni  $A, B, C$  si ha:

$$\begin{array}{lll} A = A \cap A, & A \cap B = B \cap A, & A \cap (B \cap C) = (A \cap B) \cap C, \\ A = A \cup A, & A \cup B = B \cup A, & A \cup (B \cup C) = (A \cup B) \cup C \end{array}$$

e

$$\begin{aligned} A \cap (B \cup C) &= (A \cap B) \cup (A \cap C) \\ A \cup (B \cap C) &= (A \cup B) \cap (A \cup C). \end{aligned}$$

Ad esempio, per la commutatività di  $\wedge$  abbiamo  $((x \in A) \wedge (x \in B)) \iff ((x \in B) \wedge (x \in A))$ ; ma in questa equivalenza il primo membro equivale a  $x \in A \cap B$ , il predicato da appartenenza ad  $A \cap B$ , il secondo membro equivale a  $x \in B \cap A$ , il predicato da appartenenza ad  $B \cap A$ . Quindi questi due prediciati sono equivalenti e così  $A \cap B = B \cap A$  per l'**assioma di estensionalità**. In modo analogo si ragiona per le altre uguaglianze.

Il connettivo di negazione presenta invece una difficoltà: se  $A$  è un insieme allora  $\{x \mid x \notin A\}$ , l'estensione della negazione del predicato di appartenenza ad  $A$ , non è un insieme (si veda l'[Esercizio J.3](#)). Dunque, non abbiamo a disposizione una immediata ‘traduzione’ insiemistica di  $\neg$ . Assegnati comunque due insiemi  $A$  e  $B$ , abbiamo però l'insieme  $A \setminus B = \{x \in A \mid x \notin B\}$ , e questo ci permette comunque di tradurre in formule insiemistiche tautologie sulla negazione.

*Ancora De Morgan.* Vediamo il caso delle leggi di De Morgan. Siano  $A, B$  e  $C$  insiemi, e chiamiamo  $\alpha, \beta$  e  $\gamma$  i corrispondenti prediciati di appartenenza nella variabile  $x$ :  $\alpha$  è “ $x \in A$ ”,  $\beta$  è “ $x \in B$ ” e  $\gamma$  è “ $x \in C$ ”. Allora “ $x \in A \setminus (B \cap C)$ ” equivale ad  $\alpha \wedge (\neg(\beta \wedge \gamma))$ . Ora, utilizzando la prima delle **leggi di De Morgan** e poi una delle **leggi distributive**, abbiamo:

$$(\alpha \wedge (\neg(\beta \wedge \gamma))) \iff (\alpha \wedge ((\neg\beta) \vee (\neg\gamma))) \iff ((\alpha \wedge (\neg\beta)) \vee (\alpha \wedge (\neg\gamma))).$$

Ricordando le definizioni di  $\alpha, \beta$  e  $\gamma$ , vediamo che l'ultima formula in questa catena equivale a “ $x \in (A \setminus B) \cup (A \setminus C)$ ”. Quindi “ $x \in A \setminus (B \cap C)$ ” e “ $x \in (A \setminus B) \cup (A \setminus C)$ ” sono equivalenti. Similmente, “ $x \in A \setminus (B \cup C)$ ” equivale ad  $\alpha \wedge ((\neg\beta) \wedge (\neg\gamma))$ , quindi ad  $(\alpha \wedge (\neg\beta)) \wedge (\alpha \wedge (\neg\gamma))$ , cioè a “ $x \in (A \setminus B) \cap (A \setminus C)$ ”. Abbiamo così due importanti formule, che, analogamente alle tautologie, sono note come *formule di De Morgan*: scelti comunque gli insiemi  $A, B$  e  $C$ ,

$$\begin{aligned} A \setminus (B \cap C) &= (A \setminus B) \cup (A \setminus C) \\ A \setminus (B \cup C) &= (A \setminus B) \cap (A \setminus C). \end{aligned} \tag{De Morgan}$$

*Differenza simmetrica.* Definiamo l'operazione insiemistica  $\Delta$  di *differenza simmetrica* come l'operazione corrispondente alla disgiunzione esclusiva; poniamo dunque, per ogni  $A$  e  $B$ ,

$$A \Delta B := \{x \mid (x \in A) \text{ XOR } (x \in B)\}.$$

Questa operazione ha proprietà algebriche notevoli che non sono evidenti a prima vista; per questo essa assume un ruolo centrale per lo studio di strutture importanti in informatica, come gli anelli booleani.

Le tautologie che abbiamo chiamato **esplicitazione di XOR**:

$$(p \text{ XOR } q) \iff ((p \wedge (\neg q)) \vee (q \wedge (\neg p))) \iff ((p \vee q) \wedge (\neg(p \wedge q)))$$

danno facilmente (chi legge è invitato a verificarlo):

$$A \Delta B = (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (B \cap A),$$

mentre la **commutatività** e l'**associatività** di XOR e la **distributività di  $\wedge$  rispetto a XOR** forniscono, ancora più direttamente, la commutatività e l'associatività di  $\Delta$  e la distributività di  $\cap$  rispetto a  $\Delta$ . Per ogni  $A, B, C$ , abbiamo, cioè:

$$A \Delta B = B \Delta A; \quad A \Delta (B \Delta C) = (A \Delta B) \Delta C; \quad A \cap (B \Delta C) = (A \cap B) \Delta (A \cap C).$$

<sup>34</sup>va aggiunto che  $A \cap B = \{x \mid (x \in A) \wedge (x \in B)\} = \{x \in A \mid x \in B\}$  e  $A \cup B = \{x \mid (x \in A) \vee (x \in B)\}$  sono effettivamente insiemi se lo sono  $A$  e  $B$ ; nel primo caso ciò segue dall'**assioma di separazione**, nel secondo invece da un altro, apposito assioma della teoria degli insiemi, che si chiama, guarda caso, **assioma dell'unione**.

*Parti di un fissato insieme.* La difficoltà che abbiamo incontrato con la ‘traduzione’ del connettivo di negazione scompare se limitiamo le nostre formule a parti di un prefissato insieme anziché ad insiemi arbitrari. Vediamo come. Fissiamo un insieme  $S$ . Se  $A$  è una parte di  $S$ , come abbiamo detto non esiste l’insieme degli oggetti che non appartengono ad  $A$ , ma invece esiste (per l’[assioma di separazione](#)) l’insieme degli elementi di  $S$  che non appartengono ad  $A$ ; questo insieme è  $S \setminus A$ . Possiamo dunque considerare l’operazione (unaria) di scelta del complemento in  $S$  come ‘traduzione’ insiemistica del connettivo di negazione (nel prefissato ambiente  $S$ ). Le ‘traduzioni’ degli altri connettivi logici ( $\Leftrightarrow, \Rightarrow, \wedge, \vee, \text{XOR}$ ) che avevamo a disposizione nel caso generale continuano ad essere valide senza sostanziali modifiche (ad esempio, per arbitrarie parti  $A$  e  $B$  di  $S$  vale:  $A = B \iff (\forall x \in S)(x \in A \Leftrightarrow x \in B)$ ) e possiamo, come sopra, ottenere rapidamente formule della teoria degli insiemi da tautologie.

Ad esempio, la tautologia della [doppia negazione](#) si traduce nella formula

$$S \setminus (S \setminus A) = A \quad \text{per ogni } A \subseteq S$$

(bisogna fare attenzione: questa formula vale nell’ipotesi  $A \subseteq S$ ; se  $A$  ed  $S$  sono insiemi arbitrari si ha  $S \setminus (S \setminus A) = S \cap A$ ; vedi l’[Esercizio J.4](#)).

Cosa ricaviamo di nuovo dalle tautologie sulla implicazione? Se  $A$  e  $B$  sono parti di  $S$ , sappiamo che  $A \subseteq B$  equivale a  $(\forall x \in S)(x \in A \Rightarrow x \in B)$ . La tautologia della [implicazione come disgiunzione](#) e la [legge di contrapposizione](#) (ricordiamo:  $(p \Rightarrow q) \iff ((\neg p) \vee q) \iff ((\neg q) \Rightarrow (\neg p))$ ) mostrano che “ $x \in A \Rightarrow x \in B$ ” equivale da una parte a “ $(x \notin A) \vee (x \in B)$ ”, dall’altra a “ $x \notin B \Rightarrow x \notin A$ ”. La conclusione è che, per ogni insieme  $S$  e per arbitrarie parti  $A$  e  $B$  di  $S$ ,

$$A \subseteq B \iff S = (S \setminus A) \cup B \iff S \setminus B \subseteq S \setminus A.$$

### Osservazioni ed Esercizi.

**J.1.** Quello descritto nell’ultima sezione di queste note è solo uno dei tanti metodi per provare formule insiemistiche. Esistono tanti altri metodi, ad esempio quello dei diagrammi di Euler-Venn; a seconda dei casi (e dei gusti) può essere conveniente usare i metodi descritti qui o uno degli altri.

Si incoraggia chi legge a verificare le formule dimostrate in quest’ultima sezione *anche* utilizzando i diagrammi di Euler-Venn.

**J.2.** Siano  $\varphi$  e  $\psi$  due predicati unari nella variabile  $x$ . Nell’ipotesi che le loro estensioni  $A_\varphi = \{x \mid \varphi\}$  e  $A_\psi = \{x \mid \psi\}$  siano insiemi, si ha:  $A_\varphi = A_\psi \iff (\forall x(\varphi \Leftrightarrow \psi))$  e  $A_\varphi \subseteq A_\psi \iff (\forall x(\varphi \Rightarrow \psi))$ .

**J.3.** Dimostrare che, come detto [nel testo](#), se  $A$  è un insieme allora non esiste l’insieme degli oggetti che non appartengono ad  $A$ . Suggerimento: se esistesse questo insieme, allora la sua unione con  $A$  sarebbe …

**J.4.** Provare, per *arbitrari* insiemi  $S$  ed  $A$  che  $S \setminus (S \setminus A) = S \cap A$ . Suggerimento: una volta che si sia osservato che  $S \cap A \subseteq S$  e  $S \setminus A = S \setminus (S \cap A)$ , la dimostrazione è quasi completa.

**J.5.** Anche il [principio del terzo escluso](#) e quello di [non contraddizione](#) hanno una traduzione insiemistica: per ogni insieme  $S$  ed ogni sua parte  $A$  si ha  $A \cup (S \setminus A) = S$  e  $A \cap (S \setminus A) = \emptyset$ . Verificarlo. Queste formule restano valide senza l’ipotesi che  $A$  sia contenuto in  $S$ ?

**J.6.** Dedurre da una delle tautologie proposte nell’[Esercizio E.8](#) la formula:

$$(\forall A, B, C)((A \subseteq B \cap C) \iff ((A \subseteq B) \wedge (A \subseteq C))).$$

La formula analoga, con  $\cup$  e  $\vee$  al posto di  $\cap$  e  $\wedge$ , non vale. Come mai? (Riguardare tutto l’[Esercizio E.8](#)).

## Alcuni esempi di tautologie

- |            |  |   |
|------------|--|---|
| <b>1.1</b> | $\neg(\neg p) \iff p$  | (legge della doppia negazione)                          |
| <b>1.2</b> | $p \vee (\neg p)$  | (legge del terzo escluso)                               |
| <b>1.3</b> | $\neg(p \wedge (\neg p))$  | (legge di non contraddizione)                           |
| <br>       |  |   |
| <b>2.1</b> | $(p \wedge (p \Rightarrow q)) \Rightarrow q$   | (legge dell'inferenza)                                  |
| <b>2.2</b> | $(p \Rightarrow q) \iff ((\neg q) \Rightarrow (\neg p))$   | (legge di contrapposizione)                             |
| <b>2.3</b> | $(p \Rightarrow q) \Rightarrow ((p \Rightarrow \neg q) \Rightarrow \neg p)$  | (riduzione all'assurdo)                                 |
| <b>2.4</b> | $(p \Rightarrow \neg p) \Rightarrow \neg p$  | (riduzione all'assurdo debole)                          |
| <b>2.5</b> | $(p \wedge (\neg p)) \Rightarrow q$  | (legge di Lewis o “ex falso quodlibet”)                 |
| <b>2.6</b> | $((p \Rightarrow q) \wedge (q \Rightarrow r)) \Rightarrow (p \Rightarrow r)$   | (transitività dell'implicazione o legge del sillogismo) |
| <b>2.7</b> | $(p \Rightarrow (q \Rightarrow r)) \iff ((p \Rightarrow q) \Rightarrow (p \Rightarrow r))$                                   | (distributività dell'implicazione)                      |
| <b>2.8</b> | $((p \Rightarrow q) \Rightarrow p) \Rightarrow p$  | (legge di Peirce)                                       |
| <br>       |  |   |
| <b>3.1</b> | $p \Rightarrow p$  | (legge dell'identità)                                   |
| <b>3.2</b> | $p \Rightarrow (q \Rightarrow p)$  | (legge dell'affermazione del conseguente)               |
| <b>3.3</b> | $(\neg p) \Rightarrow (p \Rightarrow q)$   | (legge della negazione dell'antecedente)                |
| <b>3.4</b> | $(p \Rightarrow (q \Rightarrow r)) \iff ((p \wedge q) \Rightarrow r)$  | (esportazione-importazione degli antecedenti)           |
| <b>3.5</b> | $(p \Rightarrow (q \Rightarrow r)) \iff (q \Rightarrow (p \Rightarrow r))$   | (scambio degli antecedenti)                             |
| <br>       |  |   |
| <b>4</b>   | $(p \Rightarrow q) \iff ((\neg p) \vee q)$   | (relazione fra implicazione e disgiunzione)             |
| <br>       |  |   |
| <b>5.1</b> | $(p \wedge p) \iff p$<br>$(p \vee p) \iff p$   | (leggi di idempotenza)                                  |
| <b>5.2</b> | $(p \wedge q) \iff (q \wedge p)$<br>$(p \vee q) \iff (q \vee p)$   | (leggi commutative)                                     |
| <b>5.3</b> | $((p \wedge q) \wedge r) \iff (p \wedge (q \wedge r))$<br>$((p \vee q) \vee r) \iff (p \vee (q \vee r))$                     | (leggi associative)                                     |
| <b>5.4</b> | $(p \wedge (q \vee r)) \iff ((p \wedge q) \vee (p \wedge r))$<br>$(p \vee (q \wedge r)) \iff ((p \vee q) \wedge (p \vee r))$ | (leggi distributive)                                    |
| <b>5.5</b> | $(\neg(p \wedge q)) \iff ((\neg p) \vee (\neg q))$<br>$(\neg(p \vee q)) \iff ((\neg p) \wedge (\neg q))$                     | (leggi di De Morgan)                                    |

## Tautologie e identità insiemistiche

Le formule nella metà sinistra della pagina sono tautologie (le lettere minuscole  $a, b, c$  indicano variabili proposizionali); quelle nella metà destra sono le corrispondenti formule della teoria degli insiemi, che valgono qualsiasi siano le collezioni  $A, B, C$  e la collezione  $S$  che le comprenda tutte.

Proprietà associativa:

$$\begin{array}{ll} (a \wedge b) \wedge c \iff a \wedge (b \wedge c) & (A \cap B) \cap C = A \cap (B \cap C) \\ (a \vee b) \vee c \iff a \vee (b \vee c) & (A \cup B) \cup C = A \cup (B \cup C) \end{array}$$

Proprietà commutativa:

$$\begin{array}{ll} a \wedge b \iff b \wedge a & A \cap B = B \cap A \\ a \vee b \iff b \vee a & A \cup B = B \cup A \\ (a \iff b) \iff (b \iff a) & \end{array}$$

Proprietà iterativa:

$$\begin{array}{ll} (a \wedge a) \iff a & A \cap A = A \\ (a \vee a) \iff a & A \cup A = A \end{array}$$

Proprietà distributiva:

$$\begin{array}{ll} (a \wedge (b \vee c)) \iff ((a \wedge b) \vee (a \wedge c)) & A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \\ (a \vee (b \wedge c)) \iff ((a \vee b) \wedge (a \vee c)) & A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \\ (a \Rightarrow (b \vee c)) \iff ((a \Rightarrow b) \vee (a \Rightarrow c)) & \\ (a \Rightarrow (b \wedge c)) \iff ((a \Rightarrow b) \wedge (a \Rightarrow c)) & A \subseteq (B \cap C) \iff (A \subseteq B \wedge A \subseteq C) \end{array}$$

Disgiunzione esclusiva e differenza simmetrica:

$$((a \vee b) \wedge \neg(a \wedge b)) \iff ((a \wedge \neg b) \vee (b \wedge \neg a)) \quad (A \cup B) - (A \cap B) = (A - B) \cup (B - A)$$

Doppia negazione:

$$\neg(\neg a) \iff a \quad S - (S - A) = A$$

Leggi di De Morgan:

$$\begin{array}{ll} \neg(a \wedge b) \iff ((\neg a) \vee (\neg b)) & S - (A \cap B) = (S - A) \cup (S - B) \\ \neg(a \vee b) \iff ((\neg a) \wedge (\neg b)) & S - (A \cup B) = (S - A) \cap (S - B) \end{array} \quad (*)$$

Terzo escluso e non contraddizione:

$$\begin{array}{ll} a \vee (\neg a) & S = A \cup (S - A) \\ \neg(a \wedge (\neg a)) & \emptyset = A \cap (S - A) \end{array}$$

Sull'implicazione:

$$\begin{array}{ll} (a \iff b) \iff ((a \Rightarrow b) \wedge (b \Rightarrow a)) & A = B \iff (A \subseteq B \wedge B \subseteq A) \\ (a \Rightarrow b) \iff ((\neg a) \vee b) & A \subseteq B \iff S = (S - A) \cup B \\ [contrapposizione] \quad (a \Rightarrow b) \iff ((\neg b) \Rightarrow (\neg a)) & A \subseteq B \iff S - B \subseteq S - A \\ [transitività] \quad ((a \Rightarrow b) \wedge (b \Rightarrow c)) \Rightarrow (a \Rightarrow c) & (A \subseteq B \wedge B \subseteq C) \Rightarrow A \subseteq C \\ (a \wedge (a \Rightarrow b)) \Rightarrow b & \\ \neg a \Rightarrow (a \Rightarrow b) & \\ b \Rightarrow (a \Rightarrow b) & \\ \neg(a \Rightarrow b) \iff (a \wedge (\neg b)) & \end{array}$$

---

(\*) le leggi di De Morgan, nella loro versione insiemistica, valgono con identica formulazione anche nel caso in cui  $A$  e  $B$  non siano parti di  $S$

Algebra

Corr. ed app.



## Corrispondenze

Siano  $a$  e  $b$  insiemi.

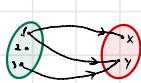
Una corrispondenza da  $a \rightarrow b$  è una terza ordinata  $(a, b, G)$  dove  $G \subseteq a \times b$ .  
 $G$  si chiama grafico della corrispondenza.

Esempio:

$$a = \{1, 2, 3\}$$

$$b = \{x, y\}$$

$$G = (a, b, \{(1, x), (1, y), (3, y)\})$$



	x	y
1	•	•
2		
3		•

$x$  e  $y$  si dicono  
corrispondenti di 1  
( $y$  lo è anche di 3).

DEF  $\forall x \in a, y \in b$  Rispetto ad  $\alpha$ :

$y$  è un corrispondente di  $x \Leftrightarrow (x, y) \in G$ .

È lecito scrivere  $x \alpha y$ .

$$G = \{(x, y) \in a \times b \mid \varphi(x, y)\}$$

Una relazione è una particolare corrispondenza da un insieme  $A$  a se stesso.

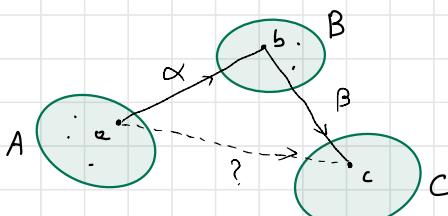
$$\text{Rel}(a) = \text{Corr}(a, a)$$

Il prodotto fra due relazioni è detto prodotto relazionale ed è così definito:

Siano  $A, B, C$  insiemi

$$\alpha \in \text{Corr}(A, B) \quad \beta \in \text{Corr}(B, C)$$

$$\alpha \beta \in \text{Corr}(A, C) : \forall a \in A, c \in C \left( a (\alpha \beta) c \Leftrightarrow \exists b \in B (a \alpha b \wedge b \beta c) \right)$$



N.B.

$$\sigma \in \text{Rel}(A)$$

$$\sigma^2 = \sigma \cdot \sigma$$

## Applicazioni

Siano  $A$  e  $B$  insiemi.

Un'applicazione  $f: A \rightarrow B$  è una  $\text{Cm}(A, B)$ :  $\forall x \in A (\exists y \in B (x \mapsto y))$

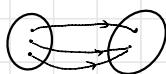


Immagine di  $x$  mediante  $f(x)$ .

Sia indicato con  $\text{Map}(A, B)$  l'insieme delle app. di dominio  $A$  e codominio  $B$ .

Allora:

$$\forall A, B, C (\forall a \in \text{Map}(A, B) \quad \forall b \in \text{Map}(B, C) \quad (ab \in \text{Map}(A, C)))$$

Il prodotto relazionale tra due applicazioni  $f: A \rightarrow B$  e  $g: B \rightarrow C$  è ancora un'applicazione (da  $A$  a  $C$ ): l'applicazione composta  $fg = g \circ f$ .

N.B. Problema della "buona definizione" di applicazione.

1.  $n \in \mathbb{N} \mapsto n+1 \in \mathbb{N}$  ✓
2.  $n \in \mathbb{N} \mapsto n-1 \in \mathbb{N}$  ✗
3.  $\{a, b\} \in P_2(\mathbb{Z}) \mapsto ab \in \mathbb{Z}$  ✓
4.  $\{a, b\} \in P_2(\mathbb{Z}) \mapsto a-b \in \mathbb{Z}$  ✗

$$P_k(A) = \{B \subseteq A \mid |B| = k\}$$

## Tipi di Applicazioni

- costante:  $\forall x, y \in A (f(x) = f(y))$
- identica:  $\text{id}_A: x \in A \mapsto x \in A$
- immersione:  $\forall A \quad \forall B \subseteq A (x \in B \mapsto x \in A)$   
(de  $B$  in  $A$ )  
si indica con  $B \hookrightarrow A$ .

Una **ristretta** è un'applicazione in cui si restringe il dominio.

Una **ridotta** è un'applicazione in cui si riduce il codominio (ma  $\text{im } f$  rimane inclusa).

Sia  $f: A \rightarrow B$  un'applicazione.  $\text{im } f = \{f(x) \mid x \in A\} \subseteq B$

Surettività:  $\text{im } f = B \iff \forall y \in B (\exists x \in A (y = f(x))) \iff \forall y \in B (\overline{f}(\{y\}) \neq \emptyset)$

Siano  $f: A \rightarrow B$  e  $g: B \rightarrow C$  surettive  $\Rightarrow g \circ f$  è surettiva.

Iniettività:  $\forall x, y \in A (x \neq y \Rightarrow f(x) \neq f(y)) \iff \forall y \in B (|\overline{f}(\{y\})| \leq 1)$

$$f: A \rightarrow B$$

Applicazione immagine di  $f$

$$\vec{f}: P(A) \rightarrow P(B) \quad x \in P(A) \mapsto \{f(t) \mid t \in x\} \in P(B)$$

Applicazione anti-immagine di  $f$

$$\overleftarrow{f}: P(B) \rightarrow P(A) \quad x \in P(B) \mapsto \{t \in A \mid f(t) \in x\} \in P(A)$$

Esempio:

$$f: n \in \mathbb{Z} \mapsto \ln n \in \mathbb{Z}$$

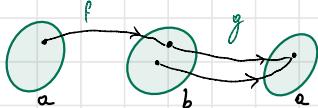
$$\vec{f}(\{-1, 0, 1\}) = \{0, 1\} = \{f(-1), f(0), f(1)\}$$

$$\vec{f}(\{2\}) = \{-2, 2\}$$

$$\vec{f}(\vec{f}(\{1\})) \neq \{1\} \quad \vec{f}(A) = \text{im } f \quad \overleftarrow{f}(B) = A = \overleftarrow{f}(\text{im } f)$$

$$\vec{f}(\vec{f}(\{x\})) \subseteq x = x \cap \text{im } f$$

$g \circ f$  biettiva  $\Rightarrow g$  suriettiva  $\wedge f$  iniettiva.



# SEZIONI E RETRAZIONI

Come sappiamo bene, la composta di due applicazioni iniettive è necessariamente iniettiva, la composta di due applicazioni suriettive è necessariamente suriettiva. Inoltre:

**1.** Siano  $f: A \rightarrow B$  e  $g: B \rightarrow C$  due applicazioni componibili.

- (i) se  $fg$  è iniettiva,  $f$  è iniettiva;
- (ii) se  $fg$  è suriettiva,  $g$  è suriettiva;
- (iii) se  $fg$  è biettiva,  $f$  è iniettiva e  $g$  è suriettiva.

N.B.

$$\begin{aligned} fg &= g \circ f \\ x^f &= f(x) \end{aligned}$$

*Dimostrazione.* Iniziamo col provare (i). Sia  $fg$  iniettiva; per ogni  $x, y \in A$ , se  $x^f = y^f$  allora  $x^{fg} = (x^f)^g = (y^f)^g = y^{fg}$ . Ma allora, poiché  $fg$  è iniettiva, si ha  $x = y$ . Abbiamo così dimostrato:  $(\forall x, y \in A)(x^f = y^f \Rightarrow x = y)$ , cioè:  $f$  è iniettiva, come richiesto dalla (i).

Proviamo (ii). Sia  $fg$  suriettiva. Per ogni  $c \in C$  esiste  $a \in A$  tale che  $c = a^{fg}$ . Posto  $b = a^f$  abbiamo dunque  $c = b^g$ . È così provato che ogni elemento di  $C$  è immagine mediante  $g$  di un elemento di  $B$ , ovvero che  $g$  è suriettiva. Anche la (ii) è così dimostrata; la (iii) è immediata conseguenza delle prime due.  $\square$

Sia  $f: A \rightarrow B$  un'applicazione. Per definizione, un'applicazione  $g: B \rightarrow A$  si dice:

sezione	di $f$	se $gf = \text{id}_B$
retrazione	di $f$	se $fg = \text{id}_A$
inversa	di $f$	se $g$ è sia una sezione che una retrazione di $f$ .

Come è evidente dalle definizioni, dire che un'applicazione  $g$  è una sezione di un'applicazione  $f$  equivale a dire che  $f$  è una retrazione di  $g$ . Una semplice dimostrazione algebrica mostra che se un'applicazione ha sia una sezione che una retrazione, queste devono coincidere. Ciò prova, in particolare, l'unicità della (eventuale) inversa di un'applicazione.

**2.** Sia  $f: A \rightarrow B$  un'applicazione e supponiamo che  $f$  abbia una sezione  $g$  ed una retrazione  $h$ . Allora  $g = h$ , in particolare  $g$  è un'inversa di  $f$ . Inoltre  $g$  è sia l'unica sezione che l'unica retrazione di  $f$ .

*Dimostrazione.* Poiché  $g$  è una sezione di  $f$  si ha  $gf = \text{id}_B$ , poiché  $h$  è una retrazione di  $f$  si ha  $fh = \text{id}_A$ . Allora  $g = g \text{id}_A = g(fh) = (gf)h = \text{id}_B h = h$ . Dunque,  $g$  è sia una sezione che una retrazione di  $f$ , quindi ne è una inversa. Se  $g_1$  è una qualsiasi sezione di  $f$ , applicando a  $g_1$  e  $h$  la prima parte dell'enunciato, appena dimostrata, si ottiene  $g_1 = h$ , quindi  $g_1 = g$ . Per lo stesso motivo, se  $h_1$  è una retrazione di  $f$  si deve avere  $g = h_1$ . Ciò prova l'unicità di  $g$  come sezione e come retrazione di  $f$ .  $\square$

Dunque, un'applicazione è *invertibile* (cioè ha un'inversa) se e solo se ha sia una sezione che una retrazione. Come segue subito da (2), di inversa ce n'è al massimo una:

**3.** Sia  $f$  un'applicazione invertibile. Allora  $f$  ha un'unica inversa.

L'unica inversa di un'applicazione invertibile  $f$  viene indicata come  $f^{-1}$ . Come vedremo, un'applicazione non invertibile può avere più di una sezione o più di una retrazione. Iniziamo a studiare le sezioni.

**4.** Sia  $f: A \rightarrow B$  un'applicazione. Esistono sezioni di  $f$  se e solo se  $f$  è suriettiva.

*Dimostrazione.* Se  $f$  ha una sezione  $g$ , allora  $gf = \text{id}_B$ . Poiché  $\text{id}_B$  è suriettiva (in effetti è biettiva), da (1) segue che  $f$  è suriettiva. Viceversa, se  $f$  è suriettiva possiamo definire una sezione di  $f$  nel modo che stiamo per descrivere. Per ogni  $b \in B$  l'antiimmagine  $A_b$  di  $\{b\}$  mediante  $f$  (definita come  $A_b = \{a \in A \mid a^f = b\}$ ) non è vuota; si può dunque scegliere, in modo arbitrario, un elemento  $b^* \in A_b$ . Effettuata questa scelta, poniamo  $g: b \in B \mapsto b^* \in A$ . Possiamo ora dimostrare che  $g$  è una sezione di  $f$ , cioè che  $gf = \text{id}_B$ . Sappiamo che  $gf$  e  $\text{id}_B$  hanno lo stesso dominio e lo stesso codominio ( $B$  in entrambi i casi), quindi dobbiamo solo provare che  $b^{gf} = b^{\text{id}_B}$ , cioè  $b^{gf} = b$  per ogni  $b \in B$ . Per ogni tale  $b$  abbiamo  $b^g \in A_b$ , per la definizione di  $g$ , e quindi  $b^{gf} = (b^g)^f = b$ , come si voleva. Dunque  $g$  è effettivamente una sezione di  $f$ , e così l'enunciato è provato.  $\square$

Si può approfondire l'argomentazione svolta nella parte finale dell'ultima dimostrazione per descrivere esplicitamente l'insieme di tutte le sezioni di un'assegnata applicazione suriettiva  $f: A \rightarrow B$ . Sia  $g$  un'applicazione da  $B$  ad  $A$ . Ponendo  $A_b = \{a \in A \mid a^f = b\}$  per ogni  $b \in B$ , abbiamo infatti, come osservato nella dimostrazione,

<sup>(\*)</sup>qui stiamo sorvolando su una serie difficoltà, alla quale facciamo solo un cenno. Il fatto che si possano effettuare *in simultanea* le scelte degli elementi  $b^*$ , anche, ad esempio, quando gli insiemi coinvolti siano infiniti e non abbiamo a disposizione alcun criterio di selezione per gli elementi di  $A$  non è affatto scontato. Per poter effettuare questa scelta, e costruire quindi la sezione  $g$  come stiamo qui facendo, è necessario un potente assioma della teoria degli insiemi, l'*assioma di scelta*. Senza questo assioma, non è possibile dimostrare l'enunciato (4), che si può anzi provare essere una forma equivalente dell'assioma di scelta.

che  $g$  è una sezione di  $f$  se e solo se  $(b^g)^f = b$  per ogni  $b \in B$ , ma questa condizione equivale a richiedere  $b^g \in A_b$  per ogni  $b \in B$ . Ciò mostra che le sezioni di  $f$  costruite col metodo della dimostrazione di (4) sono le uniche esistenti, quindi  $f$  ha tante sezioni quanti sono i modi di scegliere un elemento da ciascuno degli insiemi  $A_b$ .

*Esempio 1.* Poniamo  $A = \{1, 2, 3, 4, 5, 6\}$  e  $B = \{u, v, w\}$ , dove  $|B| = 3$  (cioè:  $u, v$  e  $w$  sono a due a due distinti) e sia  $f: A \rightarrow B$  definita da

$$f: \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ u & v & u & w & v & u \end{pmatrix}.$$

Chiaramente  $f$  è suriettiva. Volendo costruire sezioni di  $f$ , consideriamo le antiimmagini mediante  $f$  dei singleton degli elementi di  $B$ :  $A_u = \{1, 3, 6\}$  (antiimmagine di  $\{u\}$ ),  $A_v = \{2, 5\}$  (antiimmagine di  $\{v\}$ ) e  $A_w = \{4\}$  (antiimmagine di  $\{w\}$ ). Un'applicazione  $g: B \rightarrow A$  è una sezione di  $f$  se e solo se manda  $u$  in un elemento di  $A_u$ ,  $v$  in un elemento di  $A_v$  e  $w$  in un elemento di  $A_w$ . Quindi abbiamo a disposizione tre possibili scelte (1, 3 o 6) per  $u^g$ , due per  $v^g$  (2 o 5) ed una sola per  $w^g$ : dobbiamo porre necessariamente  $w^g = 4$ . Esistono dunque esattamente  $3 \cdot 2 \cdot 1 = 6$  sezioni di  $f$ . Esse sono le applicazioni da  $B$  ad  $A$  qui descritte:

$$\begin{pmatrix} u & v & w \\ 1 & 2 & 4 \end{pmatrix} \quad \begin{pmatrix} u & v & w \\ 1 & 5 & 4 \end{pmatrix} \quad \begin{pmatrix} u & v & w \\ 3 & 2 & 4 \end{pmatrix} \quad \begin{pmatrix} u & v & w \\ 3 & 5 & 4 \end{pmatrix} \quad \begin{pmatrix} u & v & w \\ 6 & 2 & 4 \end{pmatrix} \quad \begin{pmatrix} u & v & w \\ 6 & 5 & 4 \end{pmatrix}.$$

In generale, il numero delle sezioni di un'applicazione suriettiva  $f: A \rightarrow B$  è il prodotto delle cardinalità degli insiemi  $A_b$ , definiti come sopra, al variare di  $b$  in  $B$ . Si osservi che l'insieme  $\{A_b \mid b \in B\}$  non è altro che  $\text{coim } f$  (e  $A_b \neq A_{b'}$  se  $b$  e  $b'$  sono elementi distinti di  $B$ ). Quindi:

**5.** Il numero delle sezioni di un'applicazione suriettiva  $f$  è il prodotto  $\prod_{X \in \text{coim } f} |X|$  delle cardinalità degli elementi della coimmagine di  $f$ .

Almeno nel caso finito, il seguente enunciato si può vedere come caso particolare del precedente.

**6.** Se un'applicazione  $f: A \rightarrow B$  ha una ed una sola sezione allora essa è biettiva.

*Dimostrazione.* Affinché  $f$  abbia almeno una sezione,  $f$  deve essere suriettiva. Se poi  $f$  è suriettiva, essa ha una sola sezione se e solo se, con le notazioni adoperate sinora,  $|A_b| = 1$  per ogni  $b \in B$  (per ogni  $b \in B$  deve esistere una sola possibile scelta per l'immagine per  $b$  mediante una sezione di  $f$ ). Ma questa condizione implica che  $f$  è anche iniettiva: se  $x, y \in A$  e  $x^f = y^f$  allora  $x, y \in A_{x^f}$ ; supponendo  $|A_{x^f}| = 1$  si ha quindi  $x = y$ . È così provato che  $f$  è biettiva.  $\square$

Vedremo tra poco che, viceversa, le applicazioni biettive sono invertibili, quindi, per (2), hanno una ed una solo sezione.

Passiamo ora allo studio delle retrazioni. Solo nel caso degli insiemi non vuoti vale per le retrazioni l'analogo della (4).

**7.** Sia  $f: A \rightarrow B$  un'applicazione e supponiamo  $A \neq \emptyset$ . Esistono retrazioni di  $f$  se e solo se  $f$  è iniettiva.

*Dimostrazione.* Se  $f$  ha una retrazione  $g$ , allora  $fg = \text{id}_A$ , quindi, poiché  $\text{id}_A$  è iniettiva, anche  $f$  è iniettiva per (1). Viceversa, supponiamo  $f$  iniettiva e costruiamo una retrazione di  $f$ . Fissiamo un elemento  $u \in A$ ; lo possiamo fare perché, per ipotesi,  $A \neq \emptyset$ . Per ogni  $b \in \text{im } f$  esiste uno ed un solo  $b^* \in A$  tale che  $(b^*)^f = b$ , questo perché  $f$  è iniettiva. Definiamo  $g$  in questo modo:

$$g: b \in B \longmapsto \begin{cases} b^* & \text{se } b \in \text{im } f \\ u & \text{se } b \notin \text{im } f \end{cases} \in A.$$

Vogliamo provare che  $g$  è effettivamente una retrazione di  $f$ , cioè che  $fg = \text{id}_A$ . Ovviamente  $fg: A \rightarrow A$ , inoltre, per ogni  $a \in A$ , abbiamo  $a^f \in \text{im } f$  e  $(a^f)^* = a$ , quindi  $a^{fg} = (a^f)^g = a$ . Ciò mostra che  $fg = \text{id}_A$ , come si voleva.  $\square$

Cosa succede nel caso in cui il dominio  $A$  dell'applicazione  $f: A \rightarrow B$  considerata sia l'insieme vuoto? In questo caso  $f$  è certamente iniettiva (ogni applicazione di dominio vuoto è trivialmente iniettiva), ma ha una retrazione se e solo se  $B = \emptyset$ . Infatti, se  $B = \emptyset$  allora  $f = \text{id}_{\emptyset}$ , ed è facile concludere che  $f = f^{-1}$ , dunque  $f$  è inversa, e quindi retrazione, di se stessa; se invece  $B \neq \emptyset$  allora  $f$  non ha retrazioni per l'ottimo motivo che non esistono applicazioni da  $B$  ad  $A = \emptyset$  (come mai?). Notiamo infine una forma equivalente dell'enunciato (7): un'applicazione è iniettiva se e solo se o ha retrazioni o ha dominio vuoto.

**8.** Sia  $f$  un'applicazione. Sono equivalenti:

- (i)  $f$  è biettiva;
- (ii)  $f$  ha una sezione ed una retrazione;
- (iii)  $f$  è invertibile.

*Dimostrazione.* Che (iii) implichia (ii) è ovvio, che (ii) implichia (iii) è stato dimostrato in (2). Nel caso in cui il dominio di  $f$  non sia vuoto, l'equivalenza tra (i) e (ii) segue immediatamente da (4) e (7). Nel caso in cui il dominio sia vuoto, come abbiamo osservato subito prima di questo enunciato,  $f$  è invertibile se e solo se anche il codominio di  $f$  è vuoto; d'altra parte è chiaro che quest'ultima condizione è necessaria e sufficiente affinché  $f$  sia biettiva. Dunque, anche in questo caso (i) e (iii), e quindi (ii), sono equivalenti.  $\square$

Chiarita la situazione per quanto riguarda le inverse delle applicazioni, ritorniamo sulle retrazioni delle applicazioni iniettive allo scopo di ottenerne una descrizione esplicita (e quindi, nel caso finito, anche contarle).

**9.** Sia  $f: A \rightarrow B$  un'applicazione iniettiva e sia  $f_0: A \rightarrow \text{im } f$  la ridotta di  $f$  alla sua immagine.<sup>(b)</sup> Allora un'applicazione  $g: B \rightarrow A$  è una retrazione di  $f$  se e solo se la restrizione di  $g$  a  $\text{im } f$  è  $f_0^{-1}$ .

*Dimostrazione.* Il fatto che  $g$  sia una retrazione di  $f$  significa, per definizione, che  $fg = \text{id}_A$ . Poiché,  $fg$  ha lo stesso dominio,  $A$ , e codominio, ancora  $A$ , di  $\text{id}_A$ , ciò equivale all'essere  $a^{fg} = a^{\text{id}_A}$ , cioè  $a^{fg} = a$ , per ogni  $a \in A$ . Sia ora  $g_0$  la restrizione di  $g$  a  $\text{im } f$ . Allora, per ogni  $a \in A$ , poiché  $a^f \in \text{im } f$ , si ha  $a^{fg} = (a^f)^g = (a^f)^{g_0}$ ; inoltre  $a^{f_0} = a^f$ , quindi  $a^{fg} = a^{f_0 g_0}$ . In conclusione  $g$  è una retrazione di  $f$  se e solo se  $a^{f_0 g_0} = a$  per ogni  $a \in A$ , cioè se e solo se  $f_0 g_0 = \text{id}_A$ . Dal momento che  $f_0$  è biettiva,  $f_0$  ha una sola retrazione (come visto in (2)), cioè  $f_0^{-1}$ . Pertanto  $f_0 g_0 = \text{id}_A$  se e solo se  $g_0 = f_0^{-1}$ . In questo modo è provato che  $g$  è una retrazione di  $f$  se e solo se  $g_0 = f_0^{-1}$ , che è quanto affermato dall'enunciato.<sup>(1)</sup>  $\square$

Possiamo esprimere lo stesso enunciato anche in questa forma: con le notazioni fissate,

$$\text{le retrazioni di } f \text{ sono tutti e soli i prolungamenti di } f_0^{-1} \text{ a } B.$$

La questione ora diventa: quali e quanti sono i prolungamenti  $f_0^{-1}$  a  $B$ ? Esaminiamo il problema da un punto di vista più generale. Se  $h: S \rightarrow T$  è un'applicazione e  $X$  è un insieme contenente  $S$ , come possiamo descrivere i prolungamenti di  $h$  a  $X$ ? Intuitivamente è chiaro che un tale prolungamento si ottiene mandando ogni elemento  $x$  di  $S$  in  $x^h$  ed assegnando arbitrarie immagini (in  $T$ ) agli elementi di  $X \setminus S$ . Precisando questa idea, per ogni applicazione  $k: X \setminus S \rightarrow T$  definiamo

$$h_k: x \in X \longmapsto \begin{cases} x^h & \text{se } x \in S \\ x^k & \text{se } x \notin S \end{cases} \in T,$$

è chiaro che  $h_k$  è un prolungamento di  $h$  a  $X$ . Si ha:

**10.** Con le notazioni appena stabilite, indicando con  $\mathcal{P}$  l'insieme dei prolungamenti di  $h$  a  $X$ , l'applicazione  $p: k \in \text{Map}(X \setminus S, T) \mapsto h_k \in \mathcal{P}$  è biettiva; la sua inversa è  $q: t \in \mathcal{P} \mapsto t|_{X \setminus S} \in \text{Map}(X \setminus S, T)$ .<sup>(b)</sup>

*Dimostrazione.* Sia  $k \in \text{Map}(X \setminus S, T)$ . Allora  $k^{pq} = (k^p)^q$  è la restrizione  $X \setminus S$  di  $k^p = h_k$ . Questa restrizione è ovviamente  $k$ . Dunque  $k^{pq} = k$ , quindi  $pq = \text{id}_{\text{Map}(X \setminus S, T)}$ . Sia ora  $t \in \mathcal{P}$ . Allora  $t^{qp} = h_{t^q}$  è l'applicazione da  $X$  a  $T$  che manda ogni elemento  $x$  di  $S$  in  $x^h$  ed ogni elemento  $y$  di  $X \setminus S$  in  $y^{t^q}$ ; ora, poiché  $h$  e  $t^q$  sono restrizioni di  $t$ , si ha  $x^h = x^t$  e  $y^{t^q} = y^t$ . Quindi  $t^{qp}$  manda ogni elemento di  $X$  nella sua immagine mediante  $t$  dunque  $t^{qp} = t$ , sicché  $qp = \text{id}_{\mathcal{P}}$ . In questo modo è provato che  $q$  è l'inversa di  $p$  e quindi, per (8), che  $p$  è biettiva.  $\square$

Otteniamo a questo punto un'altra utile descrizione delle retrazioni di un'assegnata applicazione iniettiva. Questa descrizione è forse meno immediata di quella fornita in (9), ma più esplicita. Riprendiamo le notazioni di (9) per l'applicazione  $f: A \rightarrow B$  e la sua ridotta  $f_0$ , e indichiamo con  $\mathcal{R}$  l'insieme delle retrazioni di  $f$ , dunque  $\mathcal{R} = \{g \in \text{Map}(B, A) \mid fg = \text{id}_A\}$ .

**11.** Con le notazioni appena fissate, per ogni applicazione  $k: B \setminus \text{im } f \rightarrow A$  si ponga:

$$g_k: b \in B \longmapsto \begin{cases} b^{f_0^{-1}} & \text{se } b \in \text{im } f \\ b^k & \text{se } b \notin \text{im } f \end{cases} \in A.$$

Allora l'applicazione  $\theta: k \in \text{Map}(B \setminus \text{im } f, A) \mapsto g_k \in \mathcal{R}$  è biettiva.

*Dimostrazione.* La dimostrazione segue immediatamente da (9) e da (10).  $\square$

<sup>(b)</sup>quindi  $f_0$  è l'applicazione:  $a \in A \mapsto a^f \in \text{im } f$ ; chiaramente  $f_0$  è biettiva.

<sup>(1)</sup>non si faccia assolutamente confusione su questo punto: in questa dimostrazione abbiamo potuto dedurre  $g_0 = f_0^{-1}$  da  $f_0 g_0 = \text{id}_A$  soltanto perché già sapevamo che  $f_0$  è biettiva. In generale, se  $h: X \rightarrow Y$  e  $k: Y \rightarrow X$  sono applicazioni e sappiamo che  $hk = \text{id}_X$  ciò non ci basta per dedurre che  $k$  sia l'inversa di  $h$ . Otteniamo questa informazione se sappiamo anche che  $kh = \text{id}_Y$ , come richiesto dalla definizione di applicazione inversa, oppure, come nella dimostrazione, se sappiamo che almeno una tra  $h$  e  $k$  è biettiva, potendo in questo caso fare uso di (2).

<sup>(2)</sup>si ricorda che  $t|_{X \setminus S}$  indica la restrizione di  $t$  a  $X \setminus S$ .

Sintetizzando, per costruire una retrazione  $g: B \rightarrow A$  di un'applicazione iniettiva  $f: A \rightarrow B$  basta fare in modo che ogni elemento di  $\text{im } f$  venga mandato da  $g$  nella sua unica controimmagine mediante  $f$ , mentre la scelta delle immagini mediante  $g$  degli elementi di  $B \setminus \text{im } f$  è del tutto arbitraria. Può essere istruttivo ritornare alla dimostrazione di (7) ed osservare come quella dimostrazione sia una versione semplificata di quella svolta per (9). È poi importante capire bene il significato di (11). Questo enunciato mostra che le retrazioni dell'applicazione  $f$  sono tutte e sole le applicazioni  $g_\alpha$ , e che queste sono a due a due distinte, nel senso che, se  $\alpha$  e  $\beta$  sono applicazioni distinte da  $B \setminus \text{im } f$  ad  $A$ , allora  $g_\alpha \neq g_\beta$ .

**12.** Sia  $f: A \rightarrow B$  un'applicazione iniettiva tra insiemi finiti. Se  $|A| = a$  e  $|B| = b$ , il numero delle retrazioni di  $f$  è  $a^{b-a}$ .

*Dimostrazione.* Abbiamo visto in (11) che l'insieme delle retrazioni di  $f$  è equipotente a  $\text{Map}(B \setminus \text{im } f, A)$ . Poiché  $|\text{im } f| = |A| = a$  e quindi  $|B \setminus \text{im } f| = b - a$ , la cardinalità di  $\text{Map}(B \setminus \text{im } f, A)$  è  $a^{b-a}$ , il che prova l'asserto.  $\square$

Conseguenza di (11), e nel caso finito anche di (12), è:

**13.** Sia  $A$  un insieme tale che  $|A| > 1$ . Un'applicazione  $f: A \rightarrow B$  ha una ed una sola retrazione se e solo se è biettiva (in questo caso l'unica retrazione di  $f$  è  $f^{-1}$ ).

*Dimostrazione.* Se  $f$  ha una retrazione, allora  $f$  è iniettiva, come mostra (7). Se ciò accade, per (11), la retrazione è unica se e solo se  $\text{Map}(B \setminus \text{im } f, A)$  ha un solo elemento. Siccome  $A$  ha almeno due elementi, ciò accade se solo se  $B \setminus \text{im } f = \emptyset$ .<sup>(†)</sup> Dire che  $B \setminus \text{im } f = \emptyset$  significa esattamente dire che  $f$  è suriettiva, dunque biettiva. Tenendo presenti anche (3) e (8), ciò prova l'asserto.  $\square$

Si può confrontare (13) con (6). Come nel caso di (7), il risultato per le applicazioni iniettive è perfettamente analogo a quello per le applicazioni suriettive solo se un'ipotesi aggiuntiva (in questo caso  $|A| > 1$ ) garantisce che il dominio non sia ‘troppo piccolo’. Abbiamo già discusso il caso delle applicazioni con dominio vuoto, vediamo cosa accade se il dominio  $A$  ha cardinalità 1 (cioè è un singleton). Se  $|A| = 1$ , qualunque sia l'insieme  $B$ , ogni applicazione  $f$  da  $A$  a  $B$  è iniettiva; affinché una tale applicazione esista deve essere  $B \neq \emptyset$ . Purché, appunto,  $B \neq \emptyset$  si ha però  $|\text{Map}(B, A)| = 1$ , quindi  $f$  ha una ed una sola retrazione, ma  $f$  non è suriettiva se  $|B| > 1$ . Ciò spiega perché è stato necessario inserire nell'enunciato l'ipotesi che  $A$  abbia più di un elemento.

*Esempio 2.* Elenchiamo tutte le retrazioni dell'applicazione iniettiva  $f: A \rightarrow B$  definita da

$$\begin{pmatrix} u & v & w \\ 2 & 5 & 3 \end{pmatrix},$$

dove  $A = \{u, v, w\}$  ha cardinalità 3 e  $B = \{1, 2, 3, 4, 5\}$ . Si ha chiaramente  $\text{im } f = \{2, 3, 5\}$ . La (9) e la (11) mostrano come costruire (tutte) le retrazioni di  $f$ . L'immagine di ciascuno dei tre elementi di  $\text{im } f$  mediante una qualsiasi retrazione  $g$  di  $f$  è determinata: si deve avere  $2^g = u$ ,  $3^g = w$  e  $5^g = v$ , cioè: ciascuno dei tre elementi di  $\text{im } f$  deve essere mandato da  $g$  nell'unico elemento del quale è immagine mediante  $f$ . In altri termini, questo fa sì che la restrizione di  $g$  a  $\text{im } f$  sia l'inversa della ridotta di  $f$  alla sua immagine, come richiesto da (9). Agli altri elementi di  $B$ , cioè 1 e 4, possiamo far corrispondere arbitrari elementi di  $A$ . Nella maniera più precisa indicata in (11), consideriamo le nove applicazioni da  $B \setminus \text{im } f$  ad  $A$ :

$$\begin{cases} 1 \mapsto u \\ 4 \mapsto u \end{cases} \quad \begin{cases} 1 \mapsto u \\ 4 \mapsto v \end{cases} \quad \begin{cases} 1 \mapsto u \\ 4 \mapsto w \end{cases} \quad \begin{cases} 1 \mapsto v \\ 4 \mapsto u \end{cases} \quad \begin{cases} 1 \mapsto v \\ 4 \mapsto v \end{cases} \quad \begin{cases} 1 \mapsto v \\ 4 \mapsto w \end{cases} \quad \begin{cases} 1 \mapsto w \\ 4 \mapsto u \end{cases} \quad \begin{cases} 1 \mapsto w \\ 4 \mapsto v \end{cases} \quad \begin{cases} 1 \mapsto w \\ 4 \mapsto w \end{cases}$$

da queste, come mostra la biezione  $\theta$  di (11), si ottengono le nove retrazioni di  $f$ :

$$\begin{cases} 1 \mapsto u \\ 2 \mapsto u \\ 3 \mapsto w \\ 4 \mapsto u \\ 5 \mapsto v \end{cases} \quad \begin{cases} 1 \mapsto u \\ 2 \mapsto u \\ 3 \mapsto w \\ 4 \mapsto v \\ 5 \mapsto v \end{cases} \quad \begin{cases} 1 \mapsto u \\ 2 \mapsto u \\ 3 \mapsto w \\ 4 \mapsto w \\ 5 \mapsto v \end{cases} \quad \begin{cases} 1 \mapsto v \\ 2 \mapsto u \\ 3 \mapsto w \\ 4 \mapsto u \\ 5 \mapsto v \end{cases} \quad \begin{cases} 1 \mapsto v \\ 2 \mapsto u \\ 3 \mapsto w \\ 4 \mapsto v \\ 5 \mapsto v \end{cases} \quad \begin{cases} 1 \mapsto v \\ 2 \mapsto u \\ 3 \mapsto w \\ 4 \mapsto w \\ 5 \mapsto v \end{cases} \quad \begin{cases} 1 \mapsto w \\ 2 \mapsto u \\ 3 \mapsto w \\ 4 \mapsto u \\ 5 \mapsto v \end{cases} \quad \begin{cases} 1 \mapsto w \\ 2 \mapsto u \\ 3 \mapsto w \\ 4 \mapsto v \\ 5 \mapsto v \end{cases} \quad \begin{cases} 1 \mapsto w \\ 2 \mapsto u \\ 3 \mapsto w \\ 4 \mapsto w \\ 5 \mapsto v \end{cases}$$

*Esercizi.*

- Si indichi come costruire un numero arbitrariamente grande di retrazioni dell'immersione di  $\mathbb{N}$  in  $\mathbb{Z}$  e di sezioni dell'applicazione  $n \in \mathbb{Z} \mapsto |n| \in \mathbb{N}$ .
- Quante sono le retrazioni dell'applicazione  $n \in \{-2, -1, 0, 1, 2\} \mapsto n^2 \in \{0, 1, 2, 3, 4\}$ ?
- Provare che se  $f: A \rightarrow B$  e  $g: B \rightarrow C$  sono applicazioni suriettive componibili e se  $\bar{f}$  e  $\bar{g}$  sono sezioni di  $f$  e  $g$  rispettivamente, allora  $\bar{g}\bar{f}$  è una sezione di  $fg$ . Enunciare e provare l'analogia proposizione per le retrazioni.
- L'inversa dell'inversa di un'applicazione invertibile è l'applicazione stessa.

<sup>(†)</sup>se  $b \in B \setminus \text{im } f$  e  $u, v$  sono due elementi distinti di  $A$ , esistono almeno due applicazioni da  $B \setminus \text{im } f$  ad  $A$ : una che manda  $b$  in  $u$  e tutti gli altri elementi di  $B \setminus \text{im } f$  in  $v$ , l'altra che manda ogni elemento di  $B \setminus \text{im } f$  in  $v$ .

Algebra

Op. binarie



# Operazioni binarie

Sia  $S$  un insieme.

Un'operazione binaria in  $S$  è un'applicazione  $S \times S \rightarrow S$

Sia  $*$  un'op. binaria. Allora:

- $*$  è commutativa  $\Leftrightarrow \forall a, b \in S (a * b = b * a)$
- $*$  è associativa  $\Leftrightarrow \forall a, b, c \in S ((a * b) * c = a * (b * c))$

Sia  $t \in S$ . Allora:

- $t$  neutro a dx  $\Leftrightarrow \forall a \in S (a * t = a)$
- $t$  neutro a sx  $\Leftrightarrow \forall a \in S (t * a = a)$
- $t$  neutro  $\Leftrightarrow t$  neutro sx e dx

Se  $s$  è un neutro sx e d un neutro dx, allora  $s = d$

DIM  $d = s * d = s$  Teorema di unicità dei neutrini

Supponiamo che  $(S, *)$  abbia elemento neutro  $t$ . Allora  $\forall x, y \in S$ :

- $y$  simmetrico dx di  $x$  in  $(S, *)$   $\Leftrightarrow x * y = t$
- $y$  simmetrico sx di  $x$  in  $(S, *)$   $\Leftrightarrow y * x = t$
- $y$  simmetrico di  $x$  in  $(S, *)$   $\Leftrightarrow y$  simmetrico sx e dx di  $x$

In questo caso,  $x$  si dice simmetrizzabile.

Teorema di unicità dei simmetrici.

## Strutture Algebriche di base

- semigruppo  $(S, *)$ :  $*$  è associativa.

sotto-semigruppo: Parte chiusa non vuota in un semigruppo.

- monoide  $(S, *, t)$ : semigruppo che ammette elemento neutro  $t$ .

sotto-monoide: Parte chiusa a cui appartiene  $t$ .

- gruppo  $(S, *, t)$ : monoide in cui ogni elemento è simmetrizzabile.  
si dice abeliano se è commutativo.

sotto-gruppo: sottomonoide a cui appartiene il simmetrico di ciascun loro elemento.

oppure  $\rightsquigarrow$  una parte non vuota  $H$  di un gruppo  $\Leftrightarrow \forall x, y \in H$  il prodotto  $x * y^{-1} \in H$  ( $y^{-1}$  = simmetrico di  $y$ )

## Elementi invertibili

$(M, *, t)$  monoide

$\mathcal{U}(M) = \{x \in M \mid x \text{ è simmetrizzabile}\}$  sottomonoide :  $t \in \mathcal{U}(M)$

Gruppo degli invertibili  $(\mathcal{U}(M), *, t)$

Sia  $S$  un insieme. VS:

$\{f \in T(S) \mid f \text{ è biettiva}\} = \text{Sym}(S)$  gruppo simmetrico  
l'insieme delle permutazioni (app. biettiva da un insieme a se stesso)

Tavole di Cayley:

$(S, *)$

$*: S \times S \rightarrow S$

$S = \{x, y, z\}$

*	x	y	z
x	x*x	x*y	x*z
y	y*x	y*y	y*z
z	z*x	z*y	z*z

Dal risultato delle tavole di Cayley,  
è possibile osservare facilmente se  $*$   
è commutativa, se ci siano neutri e  
simmetrizzabili, etc.

## Cancellabilità

A livello di linguaggio informale, la parola “cancellabile” ha in algebra lo stesso significato che ha nella lingua italiana di ogni giorno: “cancellabile” significa “che si può cancellare”, intendendo con questo che chiamiamo  $a$  cancellabile quando possiamo dedurre da ogni uguaglianza della forma  $ax = ay$  (oppure  $xa = ya$ ) l’uguaglianza  $x = y$ .

Diamo una definizione più precisa. Sia  $S$  un insieme dotato di un’operazione binaria interna  $*$ , e sia  $a \in S$ . Diciamo che  $a$  è *cancellabile a sinistra* in  $(S, *)$  se e solo se si ha:

$$(\forall b, c \in S)(a * b = a * c \Rightarrow b = c).$$

Si può riformulare questa definizione in modo anche più sintetico: per ogni  $a \in S$  si considera la *traslazione sinistra* determinata da  $a$  in  $(S, *)$ , cioè l’applicazione

$$\sigma_a: x \in S \longmapsto a * x \in S;$$

dovrebbe essere chiaro che  $a$  è cancellabile a sinistra se e solo se  $\sigma_a$  è iniettiva.

Esiste ovviamente anche la nozione, analoga, di cancellabilità a destra. Fissati  $a$  e  $S$  come sopra, diciamo che  $a$  è *cancellabile a destra* in  $(S, *)$  se e solo se  $(\forall b, c \in S)(b * a = c * a \Rightarrow b = c)$ , ovvero se e solo se la *traslazione destra* determinata da  $a$  in  $(S, *)$ :

$$\delta_a: x \in S \longmapsto x * a \in S$$

è iniettiva. Si dice infine che  $a$  è *cancellabile* in  $(S, *)$  se e solo se  $a$  è cancellabile sia a sinistra che a destra in  $(S, *)$ .

Va tenuto presente che se l’operazione  $*$  è commutativa non ha senso distinguere tra cancellabilità a sinistra, cancellabilità a destra e cancellabilità: le tre proprietà sono in questo caso equivalenti.

È anche il caso di osservare esplicitamente in che modo va negata la cancellabilità: un elemento  $a$  di  $S$  non è cancellabile a sinistra in  $(S, *)$  se e solo se esistono  $b$  e  $c$  in  $S$  tali che  $a * b = a * c$  ma  $b \neq c$ ; in modo analogo si nega la cancellabilità a destra.

**Esempi.** Ogni numero intero è cancellabile in  $(\mathbb{Z}, +)$  (se  $a, b$  e  $c$  sono interi, da  $a + b = a + c$  segue senz’altro  $b = c$ ); allo stesso modo ogni intero diverso da 0 è cancellabile in  $(\mathbb{Z}, \cdot)$ , invece il numero 0 non è cancellabile in  $(\mathbb{Z}, \cdot)$ : infatti  $0 \cdot 5 = 0 \cdot 2$  ma  $5 \neq 2$ . Similmente, in  $(\mathcal{P}(\mathbb{Z}), \cup)$ ,  $\mathbb{N}$  non è cancellabile perché, ad esempio,  $\mathbb{N} \cup \{2\} = \mathbb{N} \cup \emptyset$ .

**Proposizione 1.** Sia  $(S, *, e)$  un monoide e sia  $a \in S$ . Se  $a$  è simmetrizzabile a sinistra (risp. simmetrizzabile a destra, simmetrizzabile) rispetto a  $*$ , allora  $a$  è cancellabile a sinistra (risp. cancellabile a destra, cancellabile) rispetto a  $*$ .

*Dimostrazione* — Consideriamo il caso in cui  $a$  è simmetrizzabile a sinistra. Esiste  $a' \in S$  tale che  $a' * a = e$ . Per ogni  $b, c \in S$ , se  $a * b = a * c$  abbiamo:

$$b = e * b = (a' * a) * b = a' * (a * b) = a' * (a * c) = (a' * a) * c = e * c = c.$$

In accordo con la definizione, ciò prova che  $a$  è cancellabile a sinistra rispetto a  $*$ . Per il caso della cancellabilità a destra la dimostrazione è analoga. Infine, se  $a$  è simmetrizzabile (cioè simmetrizzabile sia a sinistra che a destra), esso è cancellabile sia a sinistra che a destra (cioè cancellabile), come segue dalla simultanea applicazione dei due casi (sinistro e destro) appena considerati.  $\square$

L’enunciato precedente fornisce un modo molto semplice per giustificare il fatto che, come osservato nell’esempio precedente, tutti i numeri interi sono cancellabili in  $(\mathbb{Z}, +)$ : essi sono tutti simmetrizzabili. Invece gli interi diversi da 0, che pure sono cancellabili in  $(\mathbb{Z}, \cdot)$  non sono simmetrizzabili in questo monoide. Concludiamo dunque che, in generale, mentre la simmetrizzabilità implica la cancellabilità, l’implicazione inversa può non valere: la cancellabilità non implica necessariamente la simmetrizzabilità. Questa implicazione vale però nel caso dei monoidi (ed in un certo senso, più generalmente, per i semigruppi) *finiti*, come ora dimostreremo.

## Omomorfismi e Isomorfismi

Siano  $(S, *)$  e  $(T, \circ)$  strutture algebriche.

Un'applicazione  $f: S \rightarrow T$  è un **omomorfismo** da  $(S, *)$  a  $(T, \circ)$   $\Leftrightarrow$

$$\forall a, b \in S \quad f(a * b) = f(a) \circ f(b)$$

Un omomorfismo biettivo è detto **isomorfismo**.

Gli omomorfismi swiettivi conservano:

- 1) commutatività
- 2) associatività
- 3) elementi neutri
- 4) elementi simmetrici

Gli isomorfismi conservano **TUTTE** le proprietà algebriche, compresa la cancellabilità.

$f: X \mapsto S^X$  da  $(P(S), \cap)$  a  $(P(S), \cup)$   $\forall S$  è un isomorfismo.

Tutti i gruppi di 2 elementi sono isomorfi. In generale, se  $|S| = k$ , con  $k$  numero primo, allora le tabelle di Cayley sono uniche a meno di isomorfismi.

Sia  $S = \{1, 2, 3, 4\}$

$\sigma \in \text{Sym}(S)$ :  $1 \mapsto 2 \mapsto 3 \mapsto 4 \mapsto 1$

$\{\text{id}_S, \sigma, \sigma^2, \sigma^3\}$  si tratta di una permutazione ciclica.

Algebra

Combinatoria



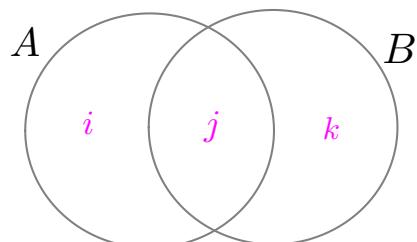
MATH

$A$  e  $B$  sono insiemi finiti

1. Quanto vale  $|A \times B|?$  risposta:  $|A \times B| = |A| \cdot |B|$

(per ogni  $x \in A$ , esistono in  $A \times B$  esattamente  $|B|$  elementi che abbiano  $x$  come prima coordinate.

2. Quanto vale  $|A \cup B|$ ?



$$i = |A \setminus B|; \quad j = |A \cap B|; \quad k = |B \setminus A|$$



$$|A| = i + j$$

$$|B| = j + k$$

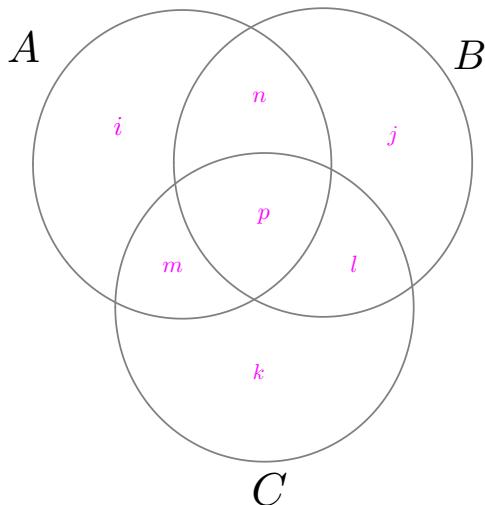
$$|A \cup B| = i + j + k = |A| + |B| - |A \cap B|$$

$$|A \cup B| + |A \setminus B| = |A| + |B|$$

$$|A| = 17 \quad |B| = 12$$

$$|A \setminus B| = 3 \Rightarrow |A \cup B| = 26$$

Per tre insiemi:



$$\begin{aligned} i &= |A \setminus (B \cup C)|; & j &= |B \setminus (A \cup C)|; & k &= |C \setminus (A \cup B)|; \\ l &= |(B \cap C) \setminus A|; & m &= |(A \cap C) \setminus B|; & n &= |(A \cap B) \setminus C|; \\ p &= |A \cap B \cap C|; \end{aligned}$$

$$\begin{aligned} |A| &= i + m + n + p; \\ |B| &= j + l + n + p; \\ |C| &= k + l + m + p \end{aligned}$$

$$|A \cup B \cup C| = \underline{i + j + k + l + m + n + p} = |A| + |B| + |C| \dots ?$$

$$\underbrace{|A| + |B| + |C|}_{|A \cup B \cup C|} - \underbrace{|A \cap B| + |B \cap C| + |A \cap C|}_{|A \cap B \cap C|}$$

principio di inclusione-esclusione

Sia  $\mathcal{S}$  un insieme finito di insiemi finiti. Posto  $n = |\mathcal{S}|$ , si ha:

$$|\bigcup \mathcal{S}| = \sum_{i=1}^n \left( (-1)^{i+1} \sum_{\mathcal{T} \in \mathcal{P}_i(\mathcal{S})} |\bigcap \mathcal{T}| \right)$$

$$\mathcal{P}_i(S) = \left\{ X \subseteq S \mid |X| = i \right\}$$

3. Quanto vale  $|\text{Map}(A, B)|$ ?

Poniamo:  $n = |A|$  e  $A = \{a_1, a_2, \dots, a_n\}$

$a_1 \mapsto ?$  ← abbiamo |B) scelte per l'immagine di  $\theta_1$   
 $a_2 \mapsto ?$  ← " " " "  
 $a_3 \mapsto ?$  ← " " " "  
 $\vdots$   
 $a_n \mapsto ?$  ← " " " "

$$|\text{Mat}_{\text{ef}}(A, B)| = |\underbrace{B \cdot B \cdot B \cdot \dots \cdot B}_n| = |B|^{|A|}$$

**Notazione:**  $B^A = \text{Map}(A, B)$

$$|B^A| = |B|^{|A|}$$

$$M_{\mathbb{R}P}(\phi, \phi) = \{\text{id}_\phi\}$$

Sie stehn im alfabeto C d k corotteri

Sia  $m \in \mathbb{N}$ . Quante parole (stringhe) di lunghezza  $m$  sull'alfabeto  $C$  esistono?

$f^m$  generalizziamo alle applicazioni

Sia  $R^M$  l'insieme delle funzioni da  $\{1, 2, \dots, n\}$  a  $C$ .  
 Si definisca tale applicazione, se essa corrisponde alle applicazioni alle quali si intende di applicare la funzione.

Osservazione:  $\forall A, B \quad f(1) = f(2) = \dots = f(m)$  / Il numero delle applicazioni costanti  $A \rightarrow B$  è  $|B|^{|A|}$ . se  $A \neq \emptyset$ , 1 se  $A = \emptyset$ .

3.1. ... e  $|\text{InjMap}(A, B)|$ ?

$a_1 \mapsto ? \leftarrow$  abbiamo  $|B|$  scelte per l'immagine di  $a_1$

$a_2 \mapsto ? \leftarrow$  "  $|B|-1$  "

$a_3 \mapsto ? \leftarrow$  "  $|B|-2$  "

⋮

$a_n \mapsto ?$

Due casi: 1) se  $|A| \leq |B|$ , procediamo in questo modo, sono al massimo  $|B| - (n - 1)$  possibili scelte per l'immagine di  $a_n$ . Allora  $|\text{InjMap}(A, B)| = |B|(|B|-1)(|B|-2) \cdots (|B|-|A|+1)$

2) se  $|A| > |B|$ , arrivati a  $a_{|B|+1}$ , scopriremo di aver utilizzato tutti gli elementi di  $B$  come possibili immagini di  $a_{|B|+1}$ , quindi non potremo procedere ulteriormente. Allora  $|\text{InjMap}(A, B)| = 0$

NB: se  $A$  e  $B$  hanno lo stesso numero (finito) di elementi, ogni applicazione iniettiva  $A \rightarrow B$  è biettiva.

$\forall n \in \mathbb{N}$

$$n! = 1 \cdot 2 \cdot 3 \cdots n = \prod_{i=1}^n i$$

(fattoriale di  $n$ ).  $0! = 1$

$$\forall n, m \in \mathbb{N}$$

$$n^m = n \cdot (n-1) \cdot (n-2) \cdots (n-m+1)$$

(fattoriale discendente).

$$m \leq n \implies n^m = \frac{n!}{(n-m)!}$$

Se  $A$  e  $B$  sono due insiemi finiti, si ha

$$|\text{InjMap}(A, B)| = |B| \underbrace{|A|}_{\begin{cases} = 0, & \text{se } |A| > |B| \\ \neq 0, & \text{se } |A| \leq |B| \end{cases}}$$

## In conclusione:

**Teorema.** Siano  $A$  e  $B$  insiemi finiti; assumiamo  $a = |A|$  e  $b = |B|$ . Allora:

(1) esistono applicazioni iniettive  $A \rightarrow B$  se e solo se  $a \leq b$ .

In questo caso, il loro numero è  $b!/(b-a)! = b^a$ .

(2) esistono applicazioni suriettive  $A \rightarrow B$  se e solo se  $a \geq b > 0$  oppure

$$a = b = 0.$$

(3) esistono applicazioni biettive  $A \rightarrow B$  se e solo se  $a = b$ .

In questo caso, il loro numero è  $a!$ .

In particolare, per ogni insieme finito  $A$ ,  $|\text{Sym}(A)| = |A|!$ .

$$\begin{aligned} 0! &= 1! = 1 \\ \text{esempi:} \\ \text{sym}(\emptyset) &= \{\text{id}_{\emptyset}\} \\ \forall x \quad \text{sym}(\{x\}) &\supset \text{id}_{\{x\}} \\ \text{se } |A|=2 & \quad |\text{sym}(A)| = 2 = 2! \\ |\text{sym}(A)| &= 2 \end{aligned}$$

DIM (1) già visto. (2): Sia  $f$  un'applicazione suriettiva da  $A$  a  $B$ . Allora  $f$  ha una sezione  $g: B \rightarrow A$ ; mentre le sezioni sono iniettive, quindi, per (1),  $b = |B| \leq |A| = a$ . Inoltre, se  $b = 0$ , necessariamente  $a = 0$ , altrimenti:

$\text{M}_{\emptyset}(A, B) = \emptyset$ . Abbiamo provato:

esistono applicazioni suriettive  $A \rightarrow B \Rightarrow a \geq b > 0$

Viceversa, se vale la condizione  $a \geq b > 0$ , allora:  
 $b = 0 \Rightarrow a = 0 \Rightarrow$  esiste un'unica applicazione  $A \rightarrow B$ ,  
 $(\text{ess. } B = \emptyset) \quad (A = \emptyset)$  cioè  $\text{id}_{\emptyset}$ , che è suriettiva.

$b \neq 0 \Rightarrow$  esiste un'applicazione iniettiva  $f: B \rightarrow A$   
 (per (1)) e, essendo  $B \neq \emptyset$ ,  $f$  ha una  
 retrozione  $g: B \rightarrow A$  che, come tutte le  
 retrozioni, è suriettiva.  
 A questo punto (2) è provata.

(3): Se esistono apppl. biettive  $A \rightarrow B$ , allora  
 $a \leq b$  per (1) e  $a \geq b$  per (2), quindi  $a = b$ .  
 Viceversa, se  $a = b$ , allora esistono, sempre  
 per (1), applicazioni iniettive  $A \rightarrow B$  e queste  
 sono tutte biettive, per quanto esse sono  
 precedenti. Ovviamente  $a = \frac{a!}{(a-a)!} = \frac{a!}{1} = a!$  □

**Teorema.** Siano  $A$  e  $B$  insiemi finiti equipotenti. Allora, per ogni applicazione  $f: A \rightarrow B$  sono equivalenti:

- (1)  $f$  è iniettiva;
- (2)  $f$  è suriettiva;
- (3)  $f$  è biettiva.

DIM: Sappiamo, dai calcoli su  $|InjMap(A, B)|$ , che vale  $(1) \Rightarrow (3)$ .

$(3) \Rightarrow (2)$  è ovvio

$(2) \Rightarrow (1)$ : Supponiamo  $f$  suriettiva. Allora  $f$  ha una sezione  $g: B \rightarrow A$ , come sappiamo  $g$  è iniettiva.

Aveva dimostrato che (1) implica (3),  
 $g$  è anche biettiva. Ma allora

$f$  è una retrozione dell'applicazione biettiva  $g$ . La l'unica retrozione d'una applicazione biettiva è la sua inversa, quindi  $f = g^{-1}$ . Allora  $f$  è biettiva.  $\square$

(in alternativa: abbiamo  $f \circ g = id_B$ ;

componendo con  $g^{-1}$  riceviamo  $f = (f \circ g) \circ g^{-1} = id_{B \circ g^{-1}} = id_B$ )

Il risultato è falso se gli insiemi sono infiniti o non hanno lo stesso numero di elementi.

$f: m \in \mathbb{N} \mapsto m+1 \in \mathbb{N}$  è iniettiva, non suriettiva

$g: m \in \mathbb{N} \mapsto \begin{cases} g & \text{se } m > 0 \\ 0 & \text{se } m = 0 \end{cases} \in \mathbb{N}$  è suriettiva, non iniettiva  
(scegli qualiasi retrozione di  $f$ )

**Definizioni.** Siano  $A$  e  $B$  insiemi.

(1)  $A$  e  $B$  sono **equipotenti** se e solo se esiste un'applicazione biettiva  $A \rightarrow B$ .

In questo caso, si scrive  $|A| = |B|$  e si dice anche che  $A$  e  $B$  hanno la stessa cardinalità.

(2) Si dice che  $A$  ha cardinalità minore o uguale a quella di  $B$  (e si scrive  $|A| \leq |B|$ ) se e solo se esiste un'applicazione iniettiva  $A \rightarrow B$ .

(3) Si dice che  $A$  ha cardinalità strettamente minore di quella di  $B$  (e si scrive  $|A| < |B|$ ) se e solo se esiste un'applicazione iniettiva  $A \rightarrow B$  ma non esistono applicazioni biettive  $A \rightarrow B$ .

**Ovviamente:** per ogni  $A$ ,  $B$  e  $C$  si ha:

$$\text{d}_A \text{ è biett.} \rightarrow |A| = |A|;$$

$$|A| = |B| \Rightarrow |B| = |A|; \quad \leftarrow f \text{ biett.} \Rightarrow f^{-1} \text{ biett.}$$

$$(|A| = |B| \wedge |B| = |C|) \Rightarrow |A| = |C|;$$

$$A \subseteq B \Rightarrow |A| \leq |B|; \quad \leftarrow \text{immersione } A \hookrightarrow B \text{ iniekt.}$$

$$|A| = |B| \Rightarrow |A| \leq |B|;$$

$$\text{simile} \rightarrow (|A| \leq |B| \wedge |B| \leq |C|) \Rightarrow |A| \leq |C|;$$

**Meno ovviamente:**

$$|\mathbb{N}^*| = |\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}| = |\mathbb{N} \times \mathbb{N}| < |\mathbb{R}| = |\mathbb{R} \times \mathbb{R}| = |\mathcal{P}(\mathbb{N})|$$

$$\forall A (|A| < |\mathcal{P}(A)|)$$

$$|A| < |\mathcal{P}(A)| < |\mathcal{P}(\mathcal{P}(A))| < |\mathcal{P}(\mathcal{P}(\mathcal{P}(A)))| \dots$$

# Funzione Caratteristica

Siano  $S$  un insieme e  $T \subseteq S$ .

$$\chi_{T,S}: a \in S \mapsto \begin{cases} 1, & \text{se } a \in T \\ 0, & \text{se } a \notin T \end{cases} \in \{0, 1\}$$

$$\alpha: T \in P(S) \mapsto \chi_{T,S} \in \text{Map}(S, \{0, 1\})$$

$$\beta: f \in \text{Map}(S, \{0, 1\}) \mapsto \overleftarrow{f}(\{1\}) \in P(S)$$

biettiva

inversa

Dimostrazione nella  
pagina seguente!

$$S = \{1, 2, 3, 4, 5, 6, 7\}$$

$$T = \{2, 4, 7\}$$

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

## Corollario

$$\forall S \quad |P(S)| = |\text{Map}(S, \{0, 1\})|$$

$$S \text{ finito} \Rightarrow |P(S)| = 2^{|S|}$$

Sia  $f: A \rightarrow B$  biettiva

Allora  $\tilde{f}: P(A) \rightarrow P(B)$  è biettiva (con inversa  $\tilde{f}^{-1}$ )

$$\forall k \in \mathbb{N}$$

$$X \in P(A) \mapsto \tilde{f}(X) \in P(B)$$

$$Y \in P(B) \mapsto \tilde{f}^{-1}(Y) \in P(A)$$

$$|P_k(S)| = ?$$

$$\text{Posto } n = |S|, \text{ allora } \binom{n}{k} := |P_k(S)|$$

coefficiente  
binomiale

$$\text{Casi banali: } \binom{n}{0} = 1 = \binom{n}{n}$$

$$\binom{n}{1} = n = \binom{n}{n-1}$$

$$k > n \Rightarrow \binom{n}{k} = 0$$

$$\sum_{k=0}^n \binom{n}{k} = 2^n$$

In generale,  
 $\binom{n}{k} = \binom{n}{n-k}$

DIM

Osserviamo: ①  $\alpha \circ \beta = \text{id}_M$   $\leftarrow$  Per comodità, poniamo  $M = M_{\text{ap}}(S, \{0, 1\})$   
 ②  $\beta \circ \alpha = \text{id}_{P(S)}$

$$\textcircled{1} \quad \forall f \in M \quad (\alpha \circ \beta(f) = \alpha(\beta(f)) = \alpha(\tilde{f}(\{1\})) = X_{\tilde{f}(\{1\}), S}$$

$$X_{\tilde{f}(\{1\}), S} : x \in S \mapsto \begin{cases} 1, & \text{se } x \in \tilde{f}(\{1\}) \\ 0, & \text{se } x \notin \tilde{f}(\{1\}) \end{cases} \in \{0, 1\}$$

$\forall x \in S$

$$x \in \tilde{f}(\{1\}) \iff f(x) = 1$$

$$x \notin \tilde{f}(\{1\}) \iff f(x) \neq 1 \iff f(x) = 0$$

$$\text{Dunque, } X_{\tilde{f}(\{1\}), S} : x \in S \mapsto \begin{cases} 1, & \text{se } f(x) = 1 \\ 0, & \text{se } f(x) = 0 \end{cases} \in \{0, 1\}$$

$$\text{Quindi: } X_{\tilde{f}(\{1\}), S} = f$$

$$\textcircled{2} \quad \forall T \in P(S) \quad (\beta \circ \alpha(T) = \beta(\alpha(T)) = \beta(X_{T, S}) = \tilde{X}_{T, S}(\{1\})).$$

$$\tilde{X}_{T, S}(\{1\}) = \{x \in S \mid X_{T, S}(x) = 1\} = T$$



# Coefficienti binomiali

Sia  $S$  un insieme. Per ogni numero naturale  $k$  si definisce  $\mathcal{P}_k(S)$  come l'insieme delle parti di  $S$  che abbiano (esattamente)  $k$  elementi, dunque:

$$\mathcal{P}_k(S) = \{X \subseteq S \mid |X| = k\}.$$

Gli elementi di  $\mathcal{P}_k(S)$  si chiamano anche  $k$ -parti di  $S$  (con una terminologia un pò vecchia ma ancora corrente, queste sono anche chiamate *combinazioni* di  $n$  oggetti di classe  $k$ ). Ad esempio, l'unica 0-parti di un insieme  $S$  è l'insieme vuoto, mentre le 1-parti di  $S$  sono i singleton degli elementi di  $S$ , quindi

$$\mathcal{P}_0(S) = \{\emptyset\} \quad \text{e} \quad \mathcal{P}_1(S) = \{\{x\} \mid x \in S\}$$

qualsiasi sia  $S$ ; se poi  $S = \{1, 2, 3, 4\}$  allora  $\mathcal{P}_2(S) = \{\{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}, \{3, 4\}\}$ .

Non è difficile verificare che se  $f: S \rightarrow T$  è un'applicazione biettiva, per ogni  $k \in \mathbb{N}$  l'applicazione

$$\alpha: X \in \mathcal{P}_k(S) \longmapsto X^{\bar{f}} \in \mathcal{P}_k(T),$$

che ad ogni  $k$ -parte di  $S$  associa la sua immagine tramite  $f$ , è ben posta ed è biettiva: se  $g$  è l'inversa di  $f$  allora l'inversa di  $\alpha$  è l'applicazione data da  $Y \in \mathcal{P}_k(T) \longmapsto Y^{\bar{g}} \in \mathcal{P}_k(S)$ . Pertanto, scelti comunque due insiemi  $S$  e  $T$  ed un numero naturale  $k$ , se  $|S| = |T|$  allora  $|\mathcal{P}_k(S)| = |\mathcal{P}_k(T)|$ .

Supponiamo ora che  $S$  sia un insieme finito e sia  $n = |S|$ . Per ogni  $k \in \mathbb{N}$  poniamo, per definizione,

$$\binom{n}{k} := |\mathcal{P}_k(S)|.$$

Il simbolo  $\binom{n}{k}$  così definito si chiama *coefficiente binomiale*. La correttezza di questa definizione va giustificata. Infatti abbiamo definito un termine (il coefficiente binomiale), che deve dipendere solo dai numeri naturali  $k$  ed  $n$ , come il numero delle  $k$ -parti di un *particolare* insieme  $S$  di  $n$  elementi. In linea di principio si potrebbe pensare che, sostituendo ad  $S$  un altro insieme con lo stesso numero  $n$  di elementi, il numero delle  $k$ -parti di questo secondo insieme possa risultare diverso da  $|\mathcal{P}_k(S)|$ ; se così fosse non avremmo correttamente definito  $\binom{n}{k}$ . Come abbiamo dimostrato sopra, però, ciò non accade: se  $T$  è un insieme e  $|T| = n = |S|$  allora  $|\mathcal{P}_k(T)| = |\mathcal{P}_k(S)|$ . In altri termini, il valore di  $|\mathcal{P}_k(S)|$  non dipende dalla particolare scelta di  $S$  tra gli insiemi con  $n$  elementi; è questo che rende accettabile la definizione data per il coefficiente binomiale.

Alcuni coefficienti binomiali sono immediati da calcolare: per ogni  $n \in \mathbb{N}$  si ha

$$\binom{n}{0} = \binom{n}{n} = 1; \quad \binom{n}{1} = n; \quad (\forall k \in \mathbb{N}) (k > n \Rightarrow \binom{n}{k} = 0).$$

Infatti, se  $S$  è un insieme di  $n$  elementi, abbiamo  $\mathcal{P}_0(S) = \{\emptyset\}$  e  $\mathcal{P}_n(S) = \{S\}$ , quindi  $\binom{n}{0} = \binom{n}{n} = 1$ ; inoltre  $\mathcal{P}_1(S)$ , l'insieme dei singleton degli elementi di  $S$ , ha tanti elementi quanto  $S$ , dunque  $n$ , quindi  $\binom{n}{1} = n$ . Infine, se  $k > n$  allora certamente  $\mathcal{P}_k(S) = \emptyset$  ( $S$  non ha parti che abbiano più elementi dello stesso  $S$ ) e quindi  $\binom{n}{k} = 0$ . Un'altra proprietà molto semplice da verificare è:

$$1. \text{ Per ogni } n \in \mathbb{N} \text{ si ha } \sum_{k=0}^n \binom{n}{k} = 2^n.$$

*Dimostrazione* — Sia  $S$  un insieme tale che  $|S| = n$ . È chiaro che  $\mathcal{P}(S)$  è unione disgiunta degli insiemi  $\mathcal{P}_k(S)$  al variare dell'intero  $k$  tra 0 e  $n$ ; in altri termini  $\{\mathcal{P}_k(S) \mid k \in \mathbb{N} \wedge k \leq n\}$  è una partizione di  $\mathcal{P}(S)$ . Pertanto  $|\mathcal{P}(S)| = \sum_{k=0}^n |\mathcal{P}_k(S)|$ ; poiché  $|\mathcal{P}(S)| = 2^n$  e, per ogni scelta di  $k$ ,  $|\mathcal{P}_k(S)| = \binom{n}{k}$ , otteniamo così l'asserto.  $\square$

Si può pensare al coefficiente binomiale  $\binom{n}{k}$  in questi termini:  $\binom{n}{k}$  è il numero di modi in cui si possono scegliere  $k$  oggetti da un insieme di  $n$  oggetti (infatti scegliere  $k$  oggetti significa in sostanza scegliere una  $k$ -parte se, come qui stiamo facendo, non consideriamo importante l'ordine in cui questi oggetti siano stati scelti). Ora, selezionare  $k$  oggetti da un insieme di  $n$  è concettualmente equivalente a sceglierne  $n - k$  da scartare (ad esempio, per essere sicuro di restare con due carte in mano se ne ho cinque, posso sceglierne due da “tenere” oppure sceglierne tre da “scartare”:  $3 = 5 - 2$ ). Dunque

dovrebbe esser facile comprendere che il coefficiente binomiale  $\binom{n}{n-k}$  coincide con  $\binom{n}{k}$ . Questa idea intuitiva è facile da formalizzare:

**2.** Siano  $n, k \in \mathbb{N}$  e supponiamo  $k \leq n$ . Allora  $\binom{n}{n-k} = \binom{n}{k}$ .

*Dimostrazione* — Fissato un insieme  $S$  con (esattamente)  $n$  elementi, consideriamo l'applicazione

$$c: X \in \mathcal{P}(S) \longmapsto S \setminus X \in \mathcal{P}(S)$$

che ad ogni parte di  $S$  associa il suo complemento in  $S$ . Poiché il complemento del complemento di una qualsiasi parte  $X$  di  $S$  è  $X$  stessa (vale a dire:  $S \setminus (S \setminus X) = X$  per ogni  $X \subseteq S$ ), è chiaro che  $c^2$  è l'applicazione identica di  $\mathcal{P}(S)$ , cioè che  $c$  è l'applicazione inversa di se stessa. Dunque  $c$  è biettiva. L'immagine di  $\mathcal{P}_k(S)$  mediante  $c$  è costituita dai complementi in  $S$  delle parti di  $S$  di cardinalità  $k$ , ma queste sono precisamente le parti di  $S$  di cardinalità  $n - k$ . Dunque, l'immagine di  $\mathcal{P}_k(S)$  mediante  $c$  è  $\mathcal{P}_{n-k}(S)$ . Pertanto l'applicazione (indotta da  $c$ , nel senso che è una restrizione di  $c$  ridotta alla sua immagine)

$$c_k: X \in \mathcal{P}_k(S) \longmapsto S \setminus X \in \mathcal{P}_{n-k}(S)$$

è anch'essa biettiva. Ciò dimostra che  $|\mathcal{P}_k(S)| = |\mathcal{P}_{n-k}(S)|$ , ovvero, in altri termini,  $\binom{n}{k} = \binom{n}{n-k}$ , come si voleva dimostrare.  $\square$

La proprietà espressa dal precedente enunciato viene talvolta chiamata proprietà di simmetria dei coefficienti binomiali. Un'altra notevolissima proprietà è quella rappresentata nel cosiddetto *triangolo di Tartaglia-Pascal*. Si tratta essenzialmente di questa formula:

**3.** Siano  $n, k \in \mathbb{N}$  e supponiamo  $k \leq n$ . Allora  $\binom{n+1}{k+1} = \binom{n}{k+1} + \binom{n}{k}$ .

*Dimostrazione* — Diamo questa dimostrazione in una versione poco formalizzata ma più facile da seguire di quanto sarebbe in una stesura più rigorosa.

Supponiamo di avere un insieme  $S$  costituito da  $n+1$  palline bianche. Ovviamente  $S \neq \emptyset$ , perché  $n+1 > 0$ , quindi possiamo selezionare una delle palline e colorarla, diciamo, di nero. Il coefficiente binomiale che vogliamo calcolare,  $\binom{n+1}{k+1}$ , è il numero delle  $(k+1)$ -parti di  $S$ . Possiamo distinguere tra due tipi di  $(k+1)$ -parti di  $S$ : quelle costituite da sole palline bianche e quelle costituite dalla pallina nera e da  $k$  palline bianche. Ovviamente  $\binom{n+1}{k+1}$  è la somma tra il numero delle parti del primo tipo ed il numero delle parti del secondo tipo. Quante sono le parti del primo tipo? Esse sono precisamente le  $(k+1)$ -parti dell'insieme delle palline bianche. Poiché il numero delle palline bianche è  $n$  (le palline erano in origine  $n+1$ , ne abbiamo colorato una di nero, restano bianche  $n = (n+1) - 1$  palline), questo numero sarà  $\binom{n}{k+1}$ . Quante sono invece le parti del secondo tipo? Ciascuna di esse si ottiene aggiungendo la pallina nera ad una  $k$ -parte dell'insieme delle palline bianche, e da ciò è facile dedurre che il numero delle parti del secondo tipo è uguale a quello delle  $k$ -parti dell'insieme delle palline bianche, dunque  $\binom{n}{k}$ . Pertanto  $\binom{n+1}{k+1}$ , che come detto è uguale alla somma tra il numero delle parti del primo tipo ed il numero delle parti del secondo tipo, è proprio  $\binom{n}{k+1} + \binom{n}{k}$ , come volevamo dimostrare.  $\square$

Vediamo come la formula appena dimostrata si visualizza nel triangolo di Tartaglia-Pascal. Questo triangolo è costruito secondo lo schema:

$$\begin{array}{ccccccc} & & \binom{0}{0} & & & & \\ & & \binom{1}{0} & \binom{1}{1} & & & \\ & & \binom{2}{0} & \binom{2}{1} & \binom{2}{2} & & \\ & & \binom{3}{0} & \binom{3}{1} & \binom{3}{2} & \binom{3}{3} & \\ & & \binom{4}{0} & \binom{4}{1} & \binom{4}{2} & \binom{4}{3} & \binom{4}{4} \\ & & & & & \dots & \end{array}$$

in cui il coefficiente binomiale  $\binom{n}{k}$  appare come  $(k+1)$ -esimo termine della riga  $(n+1)$ -esima. A parte i coefficienti che appaiono sui lati del triangolo (quelli della forma  $\binom{n}{0}$  oppure  $\binom{n}{n}$ , che sappiamo essere uguali a 1), la formula provata in (3) mostra che ciascun coefficiente è la somma dei due che gli sono sopra, cioè, nel rigo superiore, uno immediatamente a sinistra, l'altro immediatamente a destra. Ciò permette di calcolare in modo relativamente semplice i coefficienti binomiali: quelli sui lati sono già noti (sono tutti 1), quindi conosciamo già le prime due righe, il termine centrale della terza riga, cioè  $\binom{2}{1}$  lo ricaviamo come somma tra i due termini della prima, quindi  $\binom{2}{1} = 1 + 1 = 2$ , come in effetti già sapevamo. Essendo ora nota la terza riga possiamo calcolare la quarta:  $\binom{3}{1} = 1 + 2 = 3$  (i due primi termini della terza riga sono, appunto, 1 e 2) e, similmente,  $\binom{3}{2} = 2 + 1 = 3$ ; dalla quarta riga ricaviamo la quinta ... e così via. Iterando il procedimento possiamo calcolare (ricorsivamente) ciascun coefficiente binomiale a cui siamo interessati; per questo procedimento è solo necessario eseguire delle addizioni. Le prime righe sono dunque:

$$\begin{array}{ccccccc}
 & & & 1 & & & \\
 & & 1 & 1 & & & \\
 & 1 & 2 & 1 & & & \\
 1 & 3 & 3 & 1 & & & \\
 1 & 4 & 6 & 4 & 1 & & \\
 1 & 5 & 10 & 10 & 5 & 1 & \\
 1 & 6 & 15 & 20 & 15 & 6 & 1 \\
 \dots & \dots & \dots & \dots & \dots & \dots & \dots
 \end{array}$$

Ad esempio, nella settima riga il terzo termine è 15, ottenuto come  $5 + 10$ , dunque  $\binom{6}{2} = 15$ . Si può anche notare come questo triangolo sia simmetrico rispetto al suo asse verticale (ciascuno dei numeri che appare coincide con quello che occupa la posizione simmetrica rispetto a quest'asse); questa proprietà è precisamente quella espressa in (1).

Oltre al metodo offerto dal triangolo di Tartaglia-Pascal, esiste una maniera diretta di calcolare i coefficienti binomiali:

4. Siano  $n, k \in \mathbb{N}$  e supponiamo  $k \leq n$ . Allora  $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ .

*Dimostrazione* — Non è difficile dimostrare questa formula per induzione, utilizzando a questo scopo la relazione ricorsiva stabilita in (3). Possiamo però anche procedere in modo del tutto indipendente e diretto.

Sia  $S$  un insieme costituito da  $n$  elementi, e sia  $I$  un qualsiasi insieme costituito da  $k$  elementi. Sappiamo che esistono esattamente  $n^k = n!/(n-k)!$  applicazioni iniettive da  $I$  a  $S$ . Ciascuna di queste applicazioni iniettive ha per immagine una  $k$ -parte di  $S$ . Viceversa, ogni  $k$ -parte  $X$  di  $S$  è immagine di qualche applicazione iniettiva da  $I$  a  $S$ ; più precisamente possiamo osservare che le applicazioni iniettive da  $I$  ad  $S$  che hanno  $X$  come immagine sono tante quante le applicazioni biettive  $I \rightarrow X$ , quindi  $k!$ . Da questo segue subito che il numero  $n^k$  delle applicazioni iniettive da  $I$  a  $S$  è pari al prodotto tra  $\binom{n}{k}$ , il numero delle  $k$ -parti di  $S$  e  $k!$ , il numero delle applicazioni iniettive che hanno per immagine una qualsiasi prefissata  $k$ -parte di  $S$ . Dunque,

$$\frac{n!}{(n-k)!} = n^k = k! \binom{n}{k},$$

da cui segue la formula nell'enunciato.

Tutto questo dovrebbe essere sufficientemente chiaro, ma possiamo anche verificare in maggior dettaglio il fatto che l'insieme  $F_X$  delle applicazioni iniettive da  $I$  ad  $S$  che hanno  $X$  come immagine ha cardinalità uguale a quella dell'insieme  $B_X$  delle applicazioni biettive da  $I$  a  $X$ . Se  $\iota: X \hookrightarrow S$  è l'immersione di  $X$  in  $S$ , ad ogni  $f \in B_X$  possiamo associare l'applicazione  $f\iota$ , che appartiene certamente a  $F_X$ ; in questo modo definiamo l'applicazione  $f \in B_X \mapsto f\iota \in F_X$ , che si vede facilmente essere biettiva. Dunque  $|F_X| = |B_X|$ . Essendo  $|X| = |I| = k$ , sappiamo che  $|B_X| = k!$ , quindi  $|F_X| = k!$ .  $\square$

Il lettore può divertirsi a verificare che sarebbe stata possibile una impostazione differente da quella qui seguita: provare prima (4) e poi dedurre da questa formula esplicita tutte le proprietà dei coefficienti binomiali che noi abbiamo invece ricavato in precedenza.

Perché i coefficienti binomiali sono chiamati proprio così? Perché appaiono nell'espressione delle potenze di quello che tradizionalmente veniva (e viene) chiamato un binomio, un'espressione del tipo  $a + b$ . Ad esempio, siamo abituati a calcolare  $(a + b)^2 = a^2 + 2ab + b^2$ , in cui i coefficienti nel secondo termine, 1, 2 e 1, sono proprio quelli che appaiono alla terza riga del triangolo di Tartaglia-Pascal;  $(a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$ , e qui i coefficienti, 1, 3, 3 e 1 descrivono la quarta riga dello stesso triangolo. Il risultato generale che spiega questo fenomeno è la cosiddetta *formula del binomio di Newton*:

**5.** *Siano  $a$  e  $b$  due elementi di un anello, e supponiamo che valga  $ab = ba$ . Allora, per ogni intero positivo  $n$ :*

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

*Dimostrazione* — Anche in questo caso possiamo scegliere tra diverse possibili dimostrazioni, tra cui una dimostrazione per induzione che si invita a svolgere per esercizio. Diamo una dimostrazione quanto più possibile diretta. Partiamo da un esempio: supponiamo di voler calcolare  $(a + b)^3$ , cioè  $(a + b)(a + b)(a + b)$ . Utilizzando più volte la proprietà distributiva abbiamo

$$\begin{aligned} (a + b)^3 &= (a + b)(a + b)(a + b) = a(a + b)(a + b) + b(a + b)(a + b) \\ &= a(a(a + b) + b(a + b)) + b(a(a + b) + b(a + b)) \\ &= a((aa + ab) + (ba + bb)) + b((aa + ab) + (ba + bb)) \\ &= aaa + aab + aba + abb + baa + bab + bba + bbb. \end{aligned}$$

Come si vede, la terza potenza di  $a + b$  si ottiene sommando tutti i possibili prodotti con tre fattori scelti tra  $a$  e  $b$  (tenendo conto dell'ordine dei fattori). A questo punto possiamo usare il fatto che  $a$  e  $b$  commutano e quindi, ad esempio,  $aab = aba = baa = a^2b$ , per raccogliere addendi uguali e riscrivere l'uguaglianza in forma più compatta:  $(a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$ .

Passiamo ora alla dimostrazione vera e propria. Qualunque sia l'intero positivo  $n$ , segue dalla proprietà distributiva che  $(a + b)^n$  è la somma di tutti i possibili prodotti  $u_1 u_2 u_3 \dots u_n$  con  $n$  fattori ciascuno dei quali sia o  $a$  oppure  $b$ . Poiché  $ab = ba$ , ciascuno di questi prodotti si può riscrivere riordinando gli  $n$  fattori in modo da far apparire prima tutti i fattori  $a$  e poi i fattori  $b$ . Supponiamo che il numero di questi ultimi sia  $i$ ; poiché il numero totale dei fattori è  $n$ , ci saranno allora esattamente  $n - i$  fattori  $a$ ; il prodotto sarà dunque  $a^{n-i} b^i$  (ad esempio, se  $n = 5$ , il prodotto  $abaab$  si può scrivere come  $a^3 b^2$ ). A questo punto sappiamo che  $(a + b)^n$  è somma di prodotti della forma  $a^{n-i} b^i$ , ci serve solo scoprire quante volte appare ciascuno di essi in questa somma. In altri termini, fissato un intero  $i$  compreso tra 0 e  $n$ , dobbiamo calcolare in quanti modi possiamo scrivere prodotti del tipo  $u_1 u_2 u_3 \dots u_n$  (descritti come sopra) con fattori  $a$  o  $b$  in modo che il fattore  $b$  appaia esattamente  $i$  volte. Se questa proprietà è realizzata, allora l'insieme  $\{\lambda \in \{1, 2, \dots, n\} \mid u_\lambda = b\}$  è una  $i$ -parte di  $\{1, 2, \dots, n\}$ , viceversa, se  $X$  è una qualsiasi  $i$ -parte di  $\{1, 2, \dots, n\}$ , allora ponendo  $u_\lambda = b$  se  $\lambda \in X$  e  $u_\lambda = a$  se  $\lambda \in \{1, 2, \dots, n\} \setminus X$  si ha che  $u_1 u_2 u_3 \dots u_n$  è uno dei prodotti del tipo descritto in cui  $b$  appare precisamente  $i$  volte. Dunque, il numero di tali prodotti è uguale al numero delle  $i$ -parti di  $\{1, 2, \dots, n\}$ , cioè  $\binom{n}{i}$ . Questo vuol dire che si ottiene  $(a + b)^n$  come una somma in cui appare una volta  $a^n$  (essendo  $1 = \binom{n}{0}$ ),  $n = \binom{n}{1}$  volte  $a^{n-1}b$ ,  $\binom{n}{2}$  volte  $a^{n-2}b^2$ , ...,  $n = \binom{n}{n-1}$  volte  $ab^{n-1}$  e una volta  $b^n$ , perché  $1 = \binom{n}{n}$ . Questa è proprio la formula che stavamo cercando di dimostrare.  $\square$

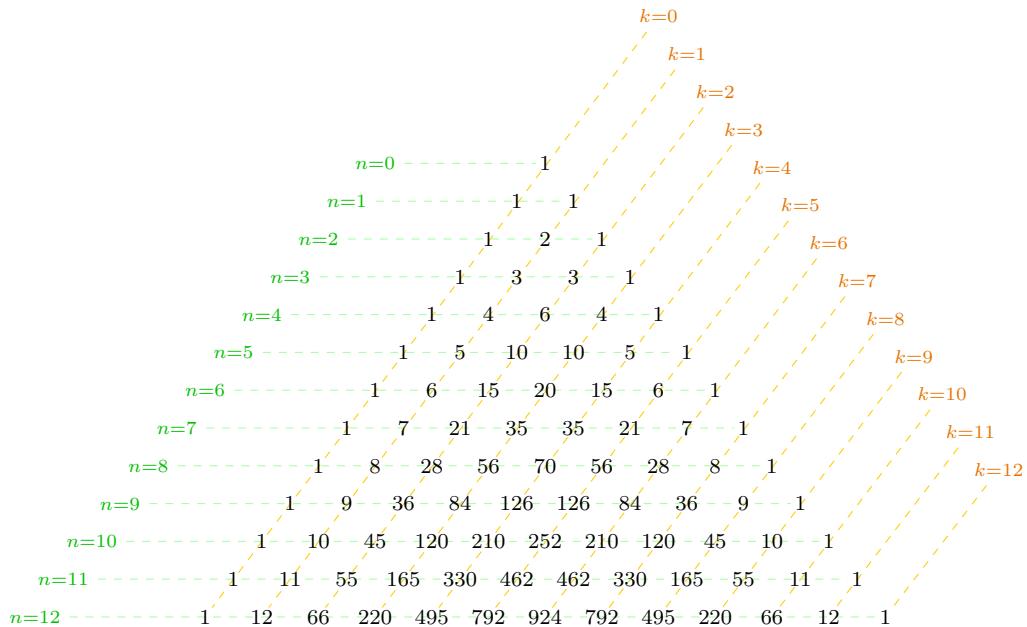
La formula di Newton vale, in particolare, per qualsiasi coppia di elementi  $a, b$  di un anello commutativo. È bene però osservare esplicitamente che essa non vale (in anelli non commutativi) nel caso in cui i due elementi  $a$  e  $b$  non commutino. Ad esempio, calcolando il quadrato di  $a + b$  potremo certamente osservare che  $(a + b)^2 = a^2 + ab + ba + b^2$ , ma se  $ab \neq ba$  questo elemento sarà certamente diverso da  $a^2 + 2ab + b^2$ .

# Il triangolo di Tartaglia-Pascal

(i coefficienti binomiali non nulli)

$$\forall n, k \in \mathbb{N} \left( \binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1} \right)$$

$$\begin{array}{ccccccccc}
& & & \binom{0}{0} & & & & & \\
& & & \binom{1}{0} & \binom{1}{1} & & & & \\
& & & \binom{2}{0} & \binom{2}{1} & \binom{2}{2} & & & \\
& & & \binom{3}{0} & \binom{3}{1} & \binom{3}{2} & \binom{3}{3} & & \\
& & & \binom{4}{0} & \binom{4}{1} & \binom{4}{2} & \binom{4}{3} & \binom{4}{4} & \\
& & & \binom{5}{0} & \binom{5}{1} & \binom{5}{2} & \binom{5}{3} & \binom{5}{4} & \binom{5}{5} \\
& & & \binom{6}{0} & \binom{6}{1} & \binom{6}{2} & \binom{6}{3} & \binom{6}{4} & \binom{6}{5} & \binom{6}{6} \\
& & & \binom{7}{0} & \binom{7}{1} & \binom{7}{2} & \binom{7}{3} & \binom{7}{4} & \binom{7}{5} & \binom{7}{6} & \binom{7}{7} \\
& & & \binom{8}{0} & \binom{8}{1} & \binom{8}{2} & \binom{8}{3} & \binom{8}{4} & \binom{8}{5} & \binom{8}{6} & \binom{8}{7} & \binom{8}{8} \\
& & & \binom{9}{0} & \binom{9}{1} & \binom{9}{2} & \binom{9}{3} & \binom{9}{4} & \binom{9}{5} & \binom{9}{6} & \binom{9}{7} & \binom{9}{8} & \binom{9}{9} \\
& & & \binom{10}{0} & \binom{10}{1} & \binom{10}{2} & \binom{10}{3} & \binom{10}{4} & \binom{10}{5} & \binom{10}{6} & \binom{10}{7} & \binom{10}{8} & \binom{10}{9} & \binom{10}{10} \\
& & & \binom{11}{0} & \binom{11}{1} & \binom{11}{2} & \binom{11}{3} & \binom{11}{4} & \binom{11}{5} & \binom{11}{6} & \binom{11}{7} & \binom{11}{8} & \binom{11}{9} & \binom{11}{10} & \binom{11}{11} \\
& & & \binom{12}{0} & \binom{12}{1} & \binom{12}{2} & \binom{12}{3} & \binom{12}{4} & \binom{12}{5} & \binom{12}{6} & \binom{12}{7} & \binom{12}{8} & \binom{12}{9} & \binom{12}{10} & \binom{12}{11} & \binom{12}{12}
\end{array}$$



# Principio d'Induzione

Sia  $p$  un predicato unario,  $b \in \mathbb{N}$   $\Rightarrow \mathbb{N}_b = \{x \in \mathbb{N} \mid x \geq b\}$ , allora:  
 $(p(b) \wedge \forall n \in \mathbb{N}_b (p(n) \Rightarrow p(n+1))) \implies (\forall n \in \mathbb{N}_b (p(n)))$

Terminologia:

$p(b)$  base d'induzione  
L'implicazione passo induttivo  
 $p(n)$  ipotesi d'induzione

Lemma

$\forall S$  insieme finito,

1. Se  $M = \{X \in P(S) \mid a \notin X\}$  e  $E = \{X \in P(S) \mid a \in X\}$ , allora  $|M| = |E|$
2.  $\forall a \in S$ ,  $|P(S)| = 2|P(S - \{a\})|$
3.  $\forall k \in \mathbb{N}$ ,  $|P_{k+1}(S)| = |P_k(S - \{a\})| + |P_k(S \cup \{a\})|$

DIM ①  $\alpha: X \in M \mapsto X \cup \{a\} \in E$  e  $\beta: X \in E \mapsto X \setminus \{a\} \in M$  sono una l'inversa dell'altra.

$$\forall X \in M (\beta \circ \alpha(X) = \beta(\alpha(X)) = \beta(X \cup \{a\}) = X \cup \{a\} \setminus \{a\} = X).$$

Dunque  $\beta \circ \alpha = \text{id}_M$

$$\forall X \in E (\alpha \circ \beta(X) = \alpha(\beta(X)) = \alpha(X \setminus \{a\}) = (X \setminus \{a\}) \cup \{a\} = X).$$

Dunque  $\alpha \circ \beta = \text{id}_E$

②  $|P(S)| = |M| + |E| = 2|M|$  per 1.

$$\text{Ma } M = \{X \in P(S) \mid a \notin X\} = P(S - \{a\}).$$

Quindi  $|P(S)| = 2|M| = 2|P(S - \{a\})|$ .

③ Poniamo  $\forall k \in \mathbb{N}$

$$M_k = M \cap P_k(S) = \{X \in P_k(S) \mid a \notin X\} = P_k(S - \{a\})$$

$$E_k = E \cap P_k(S) = \{X \in P_k(S) \mid a \in X\}$$

Siamo  $\alpha: X \in M_k \mapsto X \cup \{a\} \in E_{k+1}$  e  $\beta: X \in E_{k+1} \mapsto X \setminus \{a\} \in M_k$   
una l'inversa dell'altra (per 1).

Allora  $|E_{k+1}| = |M_k| \quad \forall k \in \mathbb{N}$

Quindi si conclude che:

$$|P_{k+1}(S)| = |M_{k+1}| + |E_{k+1}| = |M_{k+1}| + |M_k| . \quad \boxed{\text{}}$$

# PARTIZIONI ED EQUIVALENZE

GIOVANNI CUTOLO

In queste pagine vengono presentate due nozioni insiemistiche di grande importanza e strettamente collegate tra loro, quella di partizione e quella di relazione di equivalenza. Come vedremo le due nozioni sono di fatto interscambiabili.

## 1. PARTIZIONI

Sia  $A$  un insieme. Per definizione, una partizione di  $A$  è un insieme  $\mathcal{F}$  di parti non vuote di  $A$  con la proprietà che ciascun elemento di  $A$  appartenga ad uno ed un solo elemento di  $\mathcal{F}$ , vale a dire:

$$\mathcal{F} \subseteq \mathcal{P}(A) \setminus \{\emptyset\} \quad \wedge \quad \forall x \in A (\exists!y \in \mathcal{F} (x \in y)).$$

Ad esempio, se  $A$  è l'insieme  $\{1, 2, 3\}$ , allora una delle partizioni di  $A$  è l'insieme  $\{\{1\}, \{2, 3\}\}$ . Non vanno confuse tra loro le nozioni di partizione e quella di parte, che sono molto diverse tra loro benché abbiano nomi simili: quasi sempre una parte (cioè un sottoinsieme) di un insieme non ne è una partizione ed una partizione non ne è una parte. Gli elementi di una partizione vengono qualche volta chiamati anche *blocchi* della partizione.

Una semplice caratterizzazione delle partizioni è data dalla seguente proposizione.

**Proposizione 1.** *Siano  $A$  e  $\mathcal{F}$  due insiemi. Allora  $\mathcal{F}$  è una partizione di  $A$  se e solo se valgono le seguenti tre proprietà:*

- (i)  $\bigcup \mathcal{F} = A$ ;
- (ii)  $\forall b, c \in \mathcal{F} (b \neq c \Rightarrow b \cap c = \emptyset)$ ;
- (iii)  $\forall b \in \mathcal{F} (b \neq \emptyset)$ .

*Dimostrazione.* Supponiamo in primo luogo che  $\mathcal{F}$  sia una partizione di  $A$  e proviamo che valgono (i), (ii) e (iii). Per definizione di partizione,  $\emptyset \notin \mathcal{F}$ , quindi è certamente verificata la condizione (iii). Inoltre, ogni elemento di  $\mathcal{F}$  è una parte di  $A$ , quindi  $\bigcup \mathcal{F} \subseteq A$  e, viceversa, sempre per definizione di partizione, ogni elemento di  $A$  appartiene ad un elemento di  $\mathcal{F}$ , quindi  $A \subseteq \bigcup \mathcal{F}$  e possiamo concludere che vale (i)<sup>(1)</sup>. Dimostriamo (ii) ragionando per assurdo. Supponiamo che (ii) sia falsa, cioè che esistano  $b, c \in \mathcal{F}$  tali che  $b \neq c$  e  $b \cap c \neq \emptyset$ .<sup>(2)</sup> Fissati tali  $b$  e  $c$ , allora esiste  $x \in b \cap c$ ; dunque  $x$  è un elemento di  $A$  che appartiene a due elementi distinti di  $\mathcal{F}$  (a  $b$  ed a  $c$ ), in contraddizione con la definizione di partizione. Questa contraddizione mostra che così come (i) e (iii), anche (ii) è vera se  $\mathcal{F}$  è una partizione di  $A$ .

Abbiamo dimostrato che la condizione espressa nell'enunciato è necessaria affinché  $\mathcal{F}$  sia una partizione; proviamo la sufficienza. Supponiamo dunque che valgano (i), (ii) e (iii). Sia  $b$  un elemento di  $\mathcal{F}$ . Ovviamente  $b \subseteq \bigcup \mathcal{F}$ , quindi, per (i),  $b$  è una parte di  $A$  e, per (iii),  $b \neq \emptyset$ . Sia ora  $x \in A$ . Allora, per (i),  $x$  appartiene ad almeno un elemento, chiamiamolo ancora  $b$ , di  $\mathcal{F}$ . Se  $x$  appartenesse anche ad un elemento  $c$  di  $\mathcal{F}$  diverso da  $b$ , allora avremmo  $x \in b \cap c$ , quindi  $b \cap c \neq \emptyset$  e, poiché  $b \neq c$ , sarebbe contraddetta (ii). Dunque  $\mathcal{F}$  è un insieme di parti non vuote di  $A$  ed ogni elemento di  $A$  appartiene ad esattamente un elemento di  $\mathcal{F}$ , quindi  $\mathcal{F}$  è una partizione di  $A$ , come richiesto.  $\square$

In termini più sintetici (ma meno esplicativi), la proposizione ci dice che una partizione di un insieme  $A$  è un insieme di parti non vuote di  $A$ , a due a due disgiunte, la cui unione sia  $A$ .

Spesso viene usata la condizione richiesta da questa caratterizzazione come definizione della nozione di partizione. Ovviamente questo approccio è perfettamente equivalente al nostro.

Indicheremo con  $\text{Partz}(A)$  l'insieme delle partizioni dell'insieme  $A$ . Segue facilmente dalla definizione che l'unica partizione dell'insieme vuoto è l'insieme vuoto stesso (vale a dire:  $\text{Partz}(\emptyset) = \{\emptyset\}$ ). Se  $A$  è un insieme non vuoto, tra le sue partizioni ci sono certamente le cosiddette *partizioni banali*, che sono quella costituita dai singleton degli elementi di  $A$ , vale a dire  $\mathcal{P}_1(A) = \{\{x\} \mid x \in A\}$ , e quella che ha  $A$  stesso come unico elemento, vale a dire  $\{A\}$ . A titolo di esercizio, può essere utile verificare le affermazioni appena fatte e le seguenti:

- L'unica partizione dell'insieme vuoto è l'insieme vuoto stesso.
- Se  $A$  è un singleton, allora  $\{A\}$  è l'unica partizione di  $A$  (in questo caso, quindi le partizioni banali di  $A$  coincidono. Ovviamente, se  $A = \{x\}$ , allora  $\{A\} = \{\{x\}\}$ );
- Se  $|A| = 2$ , allora  $A$  ha esattamente due partizioni, quelle banali, che in questo caso non coincidono. Ad esempio, se  $A = \{1, 2\}$ , allora  $\text{Partz}(A)$  ha due elementi: la partizione  $\mathcal{P}_1(A) = \{\{1\}, \{2\}\}$  e la partizione  $\{\{A\}\}$ .

<sup>(1)</sup>ricordiamo che  $\bigcup \mathcal{F}$ , anche scritto  $\bigcup_{b \in \mathcal{F}} b$ , è l'insieme  $\{x \mid \exists b \in \mathcal{F} (x \in b)\}$  degli oggetti che appartengano ad almeno un elemento di  $\mathcal{F}$ .

<sup>(2)</sup>la negazione di (ii) è espressa da:  $\exists b, c \in \mathcal{F} (b \neq c \wedge b \cap c \neq \emptyset)$ .

- Se  $|A| = \{1, 2, 3\}$ , allora  $A$  ha esattamente cinque partizioni: le due banali,  $\mathcal{P}_1(A) = \{\{1\}, \{2\}, \{3\}\}$  e  $\{\{A\}\}$ , ed inoltre le tre partizioni  $\{\{1\}, \{2, 3\}\}$ ,  $\{\{2\}, \{1, 3\}\}$ ,  $\{\{3\}, \{2, 1\}\}$  costituite da un singleton ed un insieme di due elementi.
- Per ogni insieme  $A$  ed ogni suo sottoinsieme  $x$  tale che  $\emptyset \neq x \neq A$ , l'insieme  $\{x, A \setminus x\}$  è una partizione di  $A$ . Questo non resta vero se non si richiede la condizione  $\emptyset \neq x \neq A$ , come mai? Verificare che ogni partizione  $\mathcal{F}$  di  $A$  tale che  $|\mathcal{F}| = 2$  ha questa stessa forma, infatti se  $x$  ne è un elemento, allora l'altro elemento di  $\mathcal{F}$  è  $A \setminus x$ .

La penultima affermazione può essere giustificata (anche) facendo uso di questa osservazione davvero elementare:

- Sia  $A$  un insieme finito e sia  $\mathcal{F} \in \text{Partz}(A)$ . Allora  $|A| = \sum_{b \in \mathcal{F}} |b|$ .

Infine, dalla definizione di partizione segue che per ogni insieme  $A$  ed ogni sua partizione  $\mathcal{F}$  è ben definita l'applicazione  $\pi_{\mathcal{F}}: A \rightarrow \mathcal{F}$  che a ciascun  $x \in A$  associa l'unico  $b \in \mathcal{F}$  tale che  $x \in b$ . Chiamiamo  $\pi_{\mathcal{F}}$  la *proiezione* di  $A$  su  $\mathcal{F}$ . Abbiamo:

**Lemma 2.** *Per ogni  $A$  ed ogni  $\mathcal{F} \in \text{Partz}(A)$  la proiezione  $\pi_{\mathcal{F}}$  di  $A$  su  $\mathcal{F}$  è suriettiva.*

*Dimostrazione.* Dobbiamo provare che ogni elemento di  $\mathcal{F}$  è nell'immagine di  $\pi_{\mathcal{F}}$ . Sia  $b \in \mathcal{F}$ . Poiché, per definizione di partizione,  $b \neq \emptyset$ , esiste  $x \in b$ . Fissato un tale  $x$ , allora  $b$  è l'unico elemento di  $\mathcal{F}$  a cui  $x$  appartiene, dunque  $b = \pi_{\mathcal{F}}(x)$  e  $b \in \text{im } \pi_{\mathcal{F}}$ . Pertanto,  $\pi_{\mathcal{F}}$  è suriettiva.  $\square$

## 2. RELAZIONI DI EQUIVALENZA

Fissiamo un insieme  $A$ . Una relazione binaria  $\sim$  in  $A$  è, come ben noto, una *relazione di equivalenza* se e solo se verifica ciascuna delle tre proprietà:

- riflessiva:  $\forall x \in A (x \sim x)$ ;
- simmetrica:  $\forall x, y \in A (x \sim y \Rightarrow y \sim x)$ ;
- transitiva:  $\forall x, y, z \in A ((x \sim y \wedge y \sim z) \Rightarrow x \sim z)$ .

Indichiamo con  $\text{Eq}(A)$  l'insieme delle relazioni di equivalenza in  $A$ . Vediamo alcuni esempi; è importante che chi legge sia in grado di giustificare pienamente tutte le affermazioni che seguono:

- (1) Per ogni insieme  $A$ , come è facile verificare (farlo!) sono relazioni di equivalenza la relazione di uguaglianza in  $A$  (rispetto alla quale due qualsiasi elementi  $x$  e  $y$  di  $A$  sono in relazione se e solo  $x = y$ ) e la relazione universale in  $A$  (cioè la relazione binaria  $\tau$  in  $A$  tale che  $x \tau y$  per ogni  $x, y \in A$ ). La relazione di uguaglianza in  $A$  ha per grafico l'insieme  $\Delta_A := \{(x, x) \mid x \in A\}$ , mentre la relazione universale ha per grafico  $A \times A$ .
- (2) Sia  $\rho$  la relazione binaria in  $\mathbb{Z}$  definita da:  $\forall x, y \in \mathbb{Z} (x \rho y \iff x + y \text{ è pari})$ . Verifichiamo che  $\rho$  è di equivalenza. Per ogni  $x \in \mathbb{Z}$ , il numero  $x + x = 2x$  è pari, quindi  $x \rho x$ ; dunque vale la proprietà riflessiva. Per ogni  $x, y \in \mathbb{Z}$  abbiamo  $x \rho y \iff x + y \text{ è pari} \iff y + x \text{ è pari} \iff y \rho x$ ; vale quindi anche la proprietà simmetrica. Infine, verifichiamo la proprietà transitiva: siano  $x, y, z \in \mathbb{Z}$  ed assumiamo  $x \rho y$  e  $y \rho z$ . Allora  $x + y$  e  $y + z$  sono pari, quindi è pari la loro somma  $x + 2y + z$  e quindi anche  $x + z = (x + y) + (y + z) - 2y$ , dunque  $x \rho z$ . Pertanto  $\rho \in \text{Eq}(\mathbb{Z})$ .
- (3) Modifichiamo l'esempio precedente considerando la relazione binaria  $\rho_3$  definita in modo analogo alla precedente ma sostituendo la nozione di numero pari (cioè multiplo di 2) con quella di numero multiplo di 3. Dunque  $\rho_3$  è la relazione binaria in  $\mathbb{Z}$  definita da:  $\forall x, y \in \mathbb{Z} (x \rho_3 y \iff x + y \text{ è multiplo di 3})$ . La relazione  $\rho_3$  non è riflessiva (infatti, ad esempio,  $1 \not\rho_3 1$ ), quindi non è di equivalenza. Cambiando ancora relazione, definiamo  $\rho_3^*$ , ancora una relazione binaria in  $\mathbb{Z}$ , ponendo:  $\forall x, y \in \mathbb{Z} (x \rho_3^* y \iff (x = y \vee x \rho_3 y))$ . A differenza della precedente,  $\rho_3^*$  è riflessiva, inoltre essa è simmetrica, ma non è transitiva, ad esempio perché  $1 \rho_3^* 2$  e  $2 \rho_3^* 4$ , ma  $1 \not\rho_3^* 4$ . Dunque, neanche  $\rho_3^*$  è una relazione di equivalenza.
- (4) Sia  $\sigma$  la relazione binaria definita in  $\mathcal{P}(\mathbb{Z})$  ponendo, per ogni  $x, y \in \mathcal{P}(\mathbb{Z})$ ,  $x \sigma y \iff x \cap \mathbb{N} = y \cap \mathbb{N}$ . Per ogni  $x \in \mathcal{P}(\mathbb{Z})$  si ha  $x \cap \mathbb{N} = x \cap \mathbb{N}$ , ovvero  $x \sigma x$ , quindi  $\sigma$  è riflessiva; per ogni  $x, y \in \mathcal{P}(\mathbb{Z})$ , se  $x \sigma y$ , cioè  $x \cap \mathbb{N} = y \cap \mathbb{N}$ , allora  $y \cap \mathbb{N} = x \cap \mathbb{N}$ , ovvero  $y \sigma x$ , quindi  $\sigma$  è simmetrica; per ogni  $x, y, z \in \mathcal{P}(\mathbb{Z})$ , se  $x \sigma y$  e  $y \sigma z$ , cioè  $x \cap \mathbb{N} = y \cap \mathbb{N}$  e  $y \cap \mathbb{N} = z \cap \mathbb{N}$ , allora  $x \cap \mathbb{N} = z \cap \mathbb{N}$ , cioè  $x \sigma z$ . Abbiamo verificato anche la proprietà transitiva per  $\sigma$ , quindi  $\sigma \in \text{Eq}(\mathcal{P}(\mathbb{Z}))$ .

In modo particolare, l'ultimo degli esempi ammette un'importante generalizzazione, che andiamo ora a discutere.

Sia  $f: A \rightarrow B$  una qualsiasi applicazione di dominio l'insieme  $A$ . Si chiama *nucleo di equivalenza di  $f$*  la relazione binaria  $\mathfrak{R}_f$  definita ponendo, per ogni  $x, y \in A$ ,  $x \mathfrak{R}_f y \iff f(x) = f(y)$ . Ad esempio, la relazione di equivalenza  $\sigma$  considerata nell'ultimo degli esempi appena visti è il nucleo di equivalenza dell'applicazione  $x \in \mathcal{P}(\mathbb{Z}) \mapsto x \cap \mathbb{N} \in \mathcal{P}(\mathbb{Z})$  (ma anche, ad esempio, della sua ridotta  $x \in \mathcal{P}(\mathbb{Z}) \mapsto x \cap \mathbb{N} \in \mathcal{P}(\mathbb{N})$ ). Non solo in questo caso, ma sempre, i nuclei di equivalenza sono relazioni di equivalenza; verifichiamolo:

**Proposizione 3.** *Siano  $f: A \rightarrow B$  un'applicazione e  $\mathfrak{R}_f$  il suo nucleo di equivalenza. Allora  $\mathfrak{R}_f \in \text{Eq}(A)$ .*

*Dimostrazione.* La verifica, molto semplice, segue la falsariga dell'esempio già visto. Per ogni  $x \in A$  si ha ovviamente  $f(x) = f(x)$ , ovvero  $x \mathfrak{R}_f x$ , quindi  $\mathfrak{R}_f$  è riflessiva; per ogni  $x, y \in A$ , abbiamo  $x \mathfrak{R}_f y \iff f(x) = f(y) \iff f(y) = f(x) \iff y \mathfrak{R}_f x$ , quindi  $\mathfrak{R}_f$  è simmetrica; per ogni  $x, y, z \in A$ , se  $x \mathfrak{R}_f y$  e  $y \mathfrak{R}_f z$ , cioè  $f(x) = f(y)$  e  $f(y) = f(z)$ , allora  $f(x) = f(z)$ , cioè  $x \mathfrak{R}_f z$ ; quindi  $\mathfrak{R}_f$  è transitiva.  $\square$

In alcuni testi il nucleo di equivalenza di un'applicazione viene chiamato *equivalenza associata* (all'applicazione). La costruzione del nucleo di equivalenza fornisce immediatamente un gran numero di esempi di relazioni di equivalenza: per ottenere una relazione di equivalenza in un insieme  $A$  basta considerare una qualsiasi applicazione che abbia  $A$  come dominio ed il nucleo di equivalenza di questa. Non solo: questa costruzione permette di verificare in modo diretto che alcune relazioni binarie sono di equivalenza. Ad esempio, consideriamo la relazione binaria  $\sim$  in  $\mathbb{N}$  definita ponendo, per ogni  $x, y \in \mathbb{N}$ ,  $x \sim y \iff x^2 - 3x = y^2 - 3y$ . Si può verificare che  $\sim$  è di equivalenza procedendo, come fatto per gli esempi presentati sopra, a verificare le proprietà riflessiva, simmetrica e transitiva, ma anche, più rapidamente ed in un colpo solo, osservando che  $\sim$  è il nucleo di equivalenza dell'applicazione  $n \in \mathbb{N} \mapsto n^2 - 3n \in \mathbb{Z}$ , e quindi  $\sim \in \text{Eq}(\mathbb{N})$  per la proposizione 3.

Un punto molto importante, che discuteremo più avanti (corollario 10), è che quella dei nuclei di equivalenza non è semplicemente una costruzione che fornisce esempi di relazioni di equivalenza, ma è l'esempio più generale possibile, nel senso che le fornisce tutte: vedremo infatti che ogni relazione di equivalenza è il nucleo di equivalenza di qualche applicazione.

**Esercizio 4.** Determinare le applicazioni  $f$  tali che...

- i) ... il nucleo di equivalenza di  $f$  sia la relazione di uguaglianza nel dominio di  $f$ ;
- ii) ... il nucleo di equivalenza di  $f$  sia la relazione universale nel dominio di  $f$ .

### 3. CLASSI DI EQUIVALENZA ED INSIEME QUOTIENTE

Forse la più importante nozione legata alle relazioni di equivalenza è quella di classe di equivalenza. Siano  $A$  un insieme,  $x$  un suo elemento e  $\sim \in \text{Eq}(A)$ . La *classe di equivalenza* di  $x$  rispetto a  $\sim$  (si dice anche: “modulo  $\sim$ ”) è l'insieme

$$[x]_\sim := \{y \in A \mid y \sim x\},$$

che è ovviamente una parte di  $A$ . Osserviamo subito che, per ogni  $x \in A$ ,  $x \in [x]_\sim$ , per la proprietà riflessiva di  $\sim$  (dunque  $[x]_\sim \neq \emptyset$ ), e  $[x]_\sim = \{y \in A \mid x \sim y\}$ , dal momento che, per la proprietà simmetrica, scelti comunque  $x$  e  $y$  in  $A$  si ha  $y \sim x \iff x \sim y$ .

Con queste stesse notazioni, si chiama *insieme quoziante* (di  $A$  rispetto a  $\sim$ , o modulo  $\sim$ ) l'insieme

$$A/\sim = \{[x]_\sim \mid x \in A\}$$

di tutte le classi di equivalenza rispetto a  $\sim$  degli elementi di  $A$ .

Facciamo un esempio: se  $\rho$  è la relazione di equivalenza in  $\mathbb{Z}$  presentata all'esempio (2) di pagina 2, allora  $[0]_\rho = \{n \in \mathbb{Z} \mid n + 0 \text{ è pari}\}$ , quindi  $[0]_\rho$  è l'insieme  $P = 2\mathbb{Z}$  dei numeri interi pari. Similmente  $[1]_\rho = \{n \in \mathbb{Z} \mid n + 1 \text{ è pari}\}$  è l'insieme  $D = \mathbb{Z} \setminus 2\mathbb{Z}$  dei numeri interi dispari. Seguirà dai prossimi risultati che ci accingiamo a provare che  $\mathbb{Z}/\rho$  è una partizione di  $\mathbb{Z}$ , e di conseguenza, poiché  $\mathbb{Z} = P \cup D$ , si ha  $\mathbb{Z}/\rho = \{P, D\}$ .

Le proprietà principali delle classi di equivalenza sono raccolte nella proposizione 7; per comodità di esposizione ne verifichiamo prima un caso particolare:

**Lemma 5.** Siano  $A$  un insieme,  $\sim \in \text{Eq}(A)$  e  $x, y \in A$ . Sono allora equivalenti:

- (i)  $x \sim y$ ;
- (ii)  $x \in [y]_\sim$ ;
- (iii)  $[x]_\sim = [y]_\sim$ .

*Dimostrazione.* Per definizione di classe d'equivalenza, certamente (i)  $\iff$  (ii). Supponiamo ora che valga (i). Per ogni  $z \in [x]_\sim$  si ha, sempre per la stessa definizione,  $z \sim x$ , quindi, per la proprietà transitiva,  $z \sim y$ , cioè  $z \in [y]_\sim$ . Dunque, se vale (i), allora  $[x]_\sim \subseteq [y]_\sim$ . Ma, se vale (i) si ha anche  $y \sim x$ , per la proprietà simmetrica, quindi, scambiando i ruoli di  $x$  e di  $y$ , abbiamo anche  $[y]_\sim \subseteq [x]_\sim$ . Abbiamo così provato che (i) implica (iii). Infine, se assumiamo (iii), poiché, come sappiamo,  $x \in [x]_\sim$  per la proprietà riflessiva, concludiamo  $x \in [y]_\sim$ . Dunque (iii) implica (ii). A questo punto la dimostrazione è completa.  $\square$

**Proposizione 6.** Siano  $A$  un insieme e  $\sim \in \text{Eq}(A)$ . Allora  $A/\sim$  è una partizione di  $A$ . Inoltre, per ogni  $x \in A$ , l'unica classe di equivalenza rispetto a  $\sim$  a cui  $x$  appartenga è  $[x]_\sim$ .

*Dimostrazione.* Per ogni  $x \in A$ , sappiamo che vale  $x \in [x]_\sim$ . Se  $y \in A$  è tale che  $x \in [y]_\sim$ , allora, per il lemma precedente,  $[y]_\sim = [x]_\sim$ . Possiamo concludere che  $[x]_\sim$  è l'unica classe di equivalenza rispetto a  $\sim$  a cui  $x$  appartenga (giustificando così l'ultima parte dell'enunciato). Abbiamo anche provato che  $A/\sim$  è un insieme di parti non vuote di  $A$  con la proprietà che ogni elemento di  $A$  appartenga ad uno ed un solo elemento di  $A/\sim$ , quindi  $A/\sim \in \text{Partz}(A)$ , come richiesto dalla prima parte dell'enunciato.  $\square$

**Proposizione 7.** Siano  $A$  un insieme,  $\sim \in \text{Eq}(A)$  e  $x, y \in A$ . Sono allora equivalenti:

- (i)  $x \sim y$ ;
- (ii)  $y \sim x$ ;
- (iii)  $x \in [y]_\sim$ ;
- (iv)  $y \in [x]_\sim$ ;
- (v)  $[x]_\sim = [y]_\sim$ ;
- (vi)  $[x]_\sim \cap [y]_\sim \neq \emptyset$ .

*Dimostrazione.* (i) e (ii) sono equivalenti tra loro per la proprietà simmetrica, e, per il lemma 5, sono anche equivalenti a (iii), (iv) e (v). Se queste valgono, allora, ovviamente,  $[x]_\sim \cap [y]_\sim = [x]_\sim \neq \emptyset$ , e quindi vale (vi). Infine,  $[x]_\sim$  e  $[y]_\sim$  sono due elementi di  $A/\sim$ , che è una partizione per la proposizione 6, quindi, come segue dalla proposizione 1, se  $[x]_\sim \neq [y]_\sim$ , allora  $[x]_\sim \cap [y]_\sim = \emptyset$ . Questo vuol dire che (vi) implica (v); a questo punto la dimostrazione è completata.<sup>(3)</sup>  $\square$

Possiamo aggiungere un semplice esercizio: nelle ipotesi della proposizione 7, anche la condizione  $[x]_\sim \subseteq [y]_\sim$  equivale a  $x \sim y$ .

È utile fissare l'attenzione su questo punto stabilito nella proposizione 7: scelti comunque due elementi  $x$  e  $y$  di un insieme  $A$  sul quale sia assegnata una relazione di equivalenza  $\sim$ , si verifica necessariamente una delle due: o  $x \sim y$  e  $[x]_\sim = [y]_\sim$ , oppure  $x \not\sim y$  e  $[x]_\sim \cap [y]_\sim = \emptyset$ .

La proposizione 6 garantisce che ogni insieme quoziante è una partizione. Vale anche l'inverso: ogni partizione è l'insieme quoziante rispetto ad una relazione di equivalenza; più precisamente rispetto ad una ed una sola relazione di equivalenza. Questo è il contenuto del teorema fondamentale su partizioni e relazioni di equivalenza.

**Teorema 8.** Per ogni insieme  $A$ , l'applicazione  $\sim \in \text{Eq}(A) \mapsto A/\sim \in \text{Partz}(A)$  è biettiva.

*Dimostrazione.* Chiamiamo  $\alpha$  l'applicazione  $\sim \in \text{Eq}(A) \mapsto A/\sim \in \text{Partz}(A)$  considerata nell'enunciato. Innanzitutto, osserviamo che  $\alpha$  è ben definita per la proposizione 6. Per verificare che  $\alpha$  è iniettiva, supponiamo che  $\rho$  e  $\sigma$  siano relazioni di equivalenza in  $A$  tali che  $\alpha(\rho) = \alpha(\sigma)$ , cioè  $A/\rho = A/\sigma$ , e proviamo che di conseguenza  $\rho = \sigma$ . Per ogni  $x \in A$ , si ha ovviamente  $[x]_\rho \in A/\rho$ , ma anche  $A/\rho = A/\sigma$ , quindi  $[x]_\rho \in A/\sigma$ , vale a dire:  $[x]_\rho$  è una classe di equivalenza rispetto a  $\sigma$ . L'unico elemento di  $A/\sigma$  a cui  $x$  appartenga è  $[x]_\sigma$  (ancora per la proposizione 6), pertanto  $[x]_\sigma = [x]_\rho$ . Questo vale per ogni  $x \in A$ . Ora, per ogni  $x, y \in A$  abbiamo:

$$x \rho y \iff [x]_\rho = [y]_\rho \iff [x]_\sigma = [y]_\sigma \iff x \sigma y,$$

per via del lemma 5 e della coincidenza, appena osservata, tra le classi modulo  $\sigma$  e quelle modulo  $\rho$ . La conclusione è che  $\sigma$  e  $\rho$  coincidono. Abbiamo così provato che  $\alpha$  è iniettiva.

Proviamo ora che  $\alpha$  è suriettiva. Sia  $\mathcal{F} \in \text{Partz}(A)$ , e sia  $\pi = \pi_{\mathcal{F}}$  la proiezione di  $A$  su  $\mathcal{F}$ , cioè, ricordiamo, l'applicazione che ad ogni  $x \in A$  fa corrispondere quell'unico elemento di  $\mathcal{F}$  a cui  $x$  appartiene. Sia ora  $\sim$  il nucleo di equivalenza di  $\pi$ . Per ogni  $x \in A$  la classe  $[x]_\sim$  è l'insieme degli  $y \in A$  tali che  $\pi(y) = \pi(x)$ . Come dovrebbe essere chiaro, si ha  $\pi(y) = \pi(x)$  se e solo se  $y \in \pi(x)$ , quindi  $[x]_\sim = \pi(x)$ . Allora

$$A/\sim = \{[x]_\sim \mid x \in A\} = \{\pi(x) \mid x \in A\} = \text{im } \pi = \mathcal{F},$$

per il lemma 2. Abbiamo appena provato che  $\mathcal{F}$  è l'immagine di  $\sim$  mediante  $\alpha$ . Con questo è verificato che  $\alpha$  è suriettiva, dunque biettiva.  $\square$

La dimostrazione appena esposta fornisce anche una descrizione dell'inversa dell'applicazione  $\alpha$ . Se, nelle notazioni del teorema,  $\mathcal{F}$  è una partizione di  $A$ , allora l'immagine di  $\mathcal{F}$  mediante  $\alpha^{-1}$  è la relazione di equivalenza  $\sim$  in  $A$  descritta nella dimostrazione come nucleo di equivalenza di  $\pi$ , cioè quella definita da:  $\forall x, y \in A (x \sim y \iff \pi(x) = \pi(y))$ , o, per darne una descrizione ancora più esplicita, da:

$$\forall x, y \in A (x \sim y \iff (\exists b \in \mathcal{F} (x \in b \wedge y \in b))),$$

infatti, se  $x \in A$  e  $b \in \mathcal{F}$  sono tali che  $x \in b$ , allora  $b = \pi(x)$  (perché  $\pi(x)$  è l'unico blocco di  $\mathcal{F}$  a cui  $x$  appartenga), quindi se  $x$  e  $y$  sono due elementi di  $A$ , dire che essi appartengono ad uno stesso blocco di  $\mathcal{F}$  equivale a dire:  $\pi(x) = \pi(y)$ .

Il teorema 8 è di grande importanza: esso stabilisce che il problema di descrivere le relazioni di equivalenza in un dato insieme è essenzialmente lo stesso che quello (generalmente più facile da affrontare direttamente) dello studio delle partizioni dello stesso insieme. Per descrivere le prime basta descrivere le seconde ed usare la biezione che abbiamo chiamato  $\alpha^{-1}$  per 'tradurre' le partizioni in relazioni di equivalenza. Facciamo qualche esempio:

- (1) Sia  $A = \{n \in \mathbb{N} \mid n < 10\}$  e consideriamo la partizione  $\mathcal{F} = \{\{0, 2, 4\}, \{1\}, \{3, 9\}, \{5, 6, 7, 8\}\}$  di  $A$  (è una partizione, vero?). Qual è la relazione di equivalenza  $\sim$  di  $A$  che corrisponde ad  $\mathcal{F}$ ? In accordo con quanto appena stabilito, è quella descritta dalla proprietà che due arbitrari elementi di  $A$  siano in relazione se e solo se appartengono allo stesso blocco di  $\mathcal{F}$ . Dunque sono equivalenti tra loro 0, 2 e 4 (che però non sono equivalenti ad altri elementi di  $A$ ), 1 è equivalente solo a sé stesso;<sup>(4)</sup> sono equivalenti tra loro 3 e 9 ed infine sono equivalenti tra loro (ma non equivalenti a 1, 3 e 9) i rimanenti elementi: 5, 6, 7 e 8. Per essere ancora più esplicativi, il grafico di  $\sim$  è l'insieme

$$(\{0, 2, 4\} \times \{0, 2, 4\}) \cup \{(1, 1)\} \cup (\{3, 9\} \times \{3, 9\}) \cup (\{5, 6, 7, 8\} \times \{5, 6, 7, 8\}),$$

<sup>(3)</sup>A titolo di esercizio, o di ulteriore chiarificazione, mostriamo anche in che modo si può dedurre che (vi) implica (v) per via diretta. Assumiamo che valga (vi). Allora esiste  $z \in [x]_\sim \cap [y]_\sim$ , dunque  $z$  appartiene sia a  $[x]_\sim$  che a  $[y]_\sim$ . Ma, per la proposizione 6,  $[z]_\sim$  è l'unica classe rispetto a  $\sim$  a cui  $z$  appartenga, dunque  $[x]_\sim = [z]_\sim = [y]_\sim$  e quindi  $[x]_\sim = [y]_\sim$ .

<sup>(4)</sup>chi avesse perplessità sull'ortografia può consultare il sito dell'Accademia della Crusca, ad esempio 1, 2 o 3

cioè l'insieme

$$\{(0,0), (1,1), (2,2), (3,3), (4,4), (5,5), (6,6), (7,7), (8,8), (9,9), \\ (0,2), (2,0), (2,4), (4,2), (0,4), (4,0), (3,9), (9,3), \\ (5,6), (6,5), (5,7), (7,5), (5,8), (8,5), (6,7), (7,6), (6,8), (8,6), (7,8), (8,7)\}$$

che può essere rappresentato, in modo molto meno pesante, dalla tabella (ogni cella della tabella rappresenta un elemento di  $A \times A$ , quelle marcate da un pallino rappresentano gli elementi del grafico):

	0	1	2	3	4	5	6	7	8	9
0	•		•		•					
1		•								
2	•		•		•					
3			•							•
4	•		•		•					
5					•	•	•	•	•	
6					•	•	•	•	•	
7					•	•	•	•	•	
8					•	•	•	•	•	
9			•							•

- (2) le due relazioni di equivalenza banali in un insieme  $A \neq \emptyset$ , cioè la relazione di uguaglianza e la relazione universale, corrispondono alle due partizioni banali di  $A$ , infatti il quoziente di  $A$  rispetto alla relazione di uguaglianza è  $\mathcal{P}_1(A) = \{\{x\} \mid x \in A\}$  (ogni classe di equivalenza è un singleton), quello rispetto alla relazione universale è  $\{A\}$  (poiché tutti gli elementi di  $A$  sono in relazione tra loro,  $A$  costituisce una classe di equivalenza). Cosa cambia se  $A$  è l'insieme vuoto?
- (3) Dal teorema fondamentale (il teorema 8), segue che, per ogni insieme  $A$ , ci sono tante relazioni di equivalenza in  $A$  quante sono le partizioni di  $A$ . Ad esempio, se  $|A| = 2$ , poiché, come abbiamo visto sopra,  $A$  ha esattamente due partizioni (quelle banali),  $A$  ha anche esattamente due relazioni di equivalenza (quelle banali). Quante sono le relazioni di equivalenza in  $A$  se  $|A| < 2$ ?
- (4) Descriviamo le relazioni di equivalenza in un insieme di tre elementi:  $A = \{1, 2, 3\}$ . Per farlo ci basta elencare le partizioni di  $A$  e quindi applicare il teorema 8. Le partizioni già le conosciamo: come abbiamo visto in un esempio precedente sono in tutto cinque, le due partizioni banali e poi le tre partizioni  $F_1 = \{\{1\}, \{2, 3\}\}$ ,  $F_2 = \{\{2\}, \{1, 3\}\}$  ed  $F_3 = \{\{3\}, \{1, 2\}\}$ . Allora le relazioni di equivalenza in  $A$  saranno anch'esse cinque: le due banali (quella di uguaglianza e quella universale) e le tre relazioni di equivalenza  $\sigma_1$ ,  $\sigma_2$  e  $\sigma_3$  che corrispondono, nell'ordine, a  $F_1$ ,  $F_2$  ed  $F_3$ . Le rappresentiamo in tabella:

$\sigma_1$	1	2	3
1	•		
2		•	•
3		•	•

$\sigma_2$	1	2	3
1	•		•
2		•	
3	•		•

$\sigma_3$	1	2	3
1	•	•	
2	•	•	
3			•

- (5) Abbiamo usato il teorema 8 per elencare le relazioni di equivalenze in un insieme, vediamo ora un esempio di in cui lo stesso teorema viene usato per elencare tutte le relazioni di equivalenza (sempre in un assegnato insieme) con assegnate proprietà, traducendo questo problema in un problema espresso in termini di partizioni. Sia  $A = \{n \in \mathbb{N} \mid n < 9\}$ . Proviamo a descrivere l'insieme  $E$  di tutte le relazioni di equivalenza  $\sim$  in  $A$  tali che  $0 \sim 1 \sim 2, 3 \in [1]_\sim \subseteq [5]_\sim, 4 \in [2]_\sim \cap [6]_\sim$  e  $7 \sim 8$ . Usando la proposizione 7 (ed il piccolo esercizio che segue), possiamo verificare (farlo in dettaglio) che la condizione equivale al richiedere  $7 \sim 8$ , che gli elementi di  $X := \{0, 1, 3, 5\}$  siano equivalenti tra loro e quelli di  $Y := \{2, 4, 6\}$  siano equivalenti tra loro ma non a quelli di  $X$ . In altri termini la condizione significa precisamente questo: che  $X$  sia contenuto in una classe di equivalenza in  $A/\sim$  ed  $Y$  sia contenuto in una classe di equivalenza in  $A/\sim$  diversa da quella contenente  $X$ , ed infine che  $7 \sim 8$ . Usiamo il teorema 8 per tradurre il nostro problema in termini di partizioni di  $A$ : dobbiamo cercare le partizioni di  $A$  costituite da almeno due blocchi distinti, uno contenente  $X$  ed uno contenente  $Y$ , ed in cui 7 ed 8 appartengono allo stesso blocco; siccome  $A = X \cup Y \cup \{7, 8\}$  questo implica che i blocchi di una tale partizione sono al massimo tre. È facile riconoscere che le partizioni che soddisfano le condizioni sono tre: una con tre blocchi distinti:  $F_1 := \{X, Y, \{7, 8\}\}$ , e due con soli due blocchi:  $F_2 := \{X \cup \{7, 8\}, Y\}$  e  $F_3 := \{X, Y \cup \{7, 8\}\}$ . Quindi, per il teorema 8, l'insieme  $E$  che volevamo descrivere ha esattamente tre elementi, le tre relazioni corrispondenti a  $F_1$ ,  $F_2$  e  $F_3$ , descritte come segue. Come già avevamo osservato, rispetto a tutte e tre gli elementi di  $X$  sono equivalenti tra loro, e quelli di  $Y$  pure, ma quelli di  $X$  non sono equivalenti a quelli di  $Y$ , inoltre 7 ed 8 sono equivalenti tra loro. Rispetto alla prima relazione di equivalenza (quella corrispondente a  $F_1$ ) 7 ed 8 non sono equivalenti a nessun elemento di  $X$  o di  $Y$ , rispetto alla seconda (quella corrispondente a  $F_2$ ) 7 ed 8 sono equivalenti agli elementi di  $X$ , rispetto alla terza (quella corrispondente a  $F_3$ ) 7 ed 8 sono equivalenti agli elementi di  $Y$ .

#### 4. ANCORA SUI NUCLEI DI EQUIVALENZA

Sia  $\sim$  una relazione di equivalenza in un insieme  $A$ . Si chiama *proiezione canonica* (di  $A$  su  $A/\sim$ , oppure definita da  $\sim$ ) l'applicazione

$$\pi_\sim : x \in A \mapsto [x]_\sim \in A/\sim.$$

È evidente dalla definizione di insieme quoziante che  $\pi_\sim$  è suriettiva, ma in realtà questo è già noto dal lemma 2, dal momento che  $\pi_\sim$  non è altro che la proiezione di  $A$  su  $A/\sim$  visto come partizione di  $A$ . Da questo e dalla dimostrazione del teorema 8 si potrebbe fare seguire la proposizione seguente, di cui però forniamo una dimostrazione diretta.

**Proposizione 9.** *Sia  $\sim$  una relazione di equivalenza. Allora  $\sim$  è il nucleo di equivalenza della proiezione canonica che definisce.*

*Dimostrazione.* Continuiamo ad usare le notazioni appena introdotte; sia  $\sim \in \text{Eq}(A)$  e sia  $\pi_\sim : A \rightarrow A/\sim$  la proiezione canonica. Sia  $\rho$  il nucleo di equivalenza di  $\pi_\sim$ . Allora, per ogni  $x, y \in A$ , abbiamo:

$$x \rho y \iff \pi_\sim(x) = \pi_\sim(y) \iff [x]_\sim = [y]_\sim \iff x \sim y,$$

per il lemma 5. Dunque,  $\rho = \sim$  e l'enunciato è provato.  $\square$

È davvero importante questa conseguenza, che avevamo già annunciato:

**Corollario 10.** *Ogni relazione di equivalenza è il nucleo di equivalenza di qualche applicazione.*

In maggior dettaglio, se  $A$  è un insieme e  $\sim$  è una relazione di equivalenza in  $A$ , esiste almeno un'applicazione  $f$  di dominio  $A$  tale che  $\sim$  sia il nucleo di equivalenza di  $f$ . Non dimostriamo qui (ma non è difficile farlo) che è possibile scegliere  $f$  in modo che  $A$  sia anche il codominio di  $f$ .

Notiamo che applicazioni diverse possono avere lo stesso nucleo di equivalenza. Ad esempio, segue dall'esercizio 4 che le relazioni di equivalenza banali in un, qualsiasi, fissato insieme non vuoto sono nuclei di equivalenza di infinite applicazioni. Per fare un esempio diverso, le applicazioni  $n \in \mathbb{Z} \mapsto |n| \in \mathbb{N}$ ,  $n \in \mathbb{Z} \mapsto |n| \in \mathbb{Z}$ ,  $n \in \mathbb{Z} \mapsto n^2 \in \mathbb{Z}$  e tante altre hanno lo stesso nucleo di equivalenza: la relazione binaria in  $\mathbb{Z}$  che dichiara equivalenti due numeri interi se e solo se essi hanno lo stesso valore assoluto.

Abbiamo così che le relazioni di equivalenza si possono riguardare da almeno tre punti di vista: come particolari relazioni binarie, secondo la definizione, come concetto associato a quello di partizione (attraverso il teorema 8) e come nozione collegata a quella di applicazione: le relazioni di equivalenza sono i nuclei di equivalenza delle applicazioni. Mentre però il teorema 8 mostra che la corrispondenza tra relazioni di equivalenza e partizioni in un dato insieme è descritta da un'applicazione biettiva, e per questo studiare le partizioni dell'insieme equivale a studiarne le relazioni di equivalenza, lo stesso non vale nel caso delle applicazioni e dei corrispondenti nuclei di equivalenza.

Studiamo ora in maggior dettaglio i nuclei di equivalenza ed i corrispondenti quozianti.

**Lemma 11.** *Siano  $f : A \rightarrow B$  un'applicazione e  $\tau$  il suo nucleo di equivalenza. Allora, per ogni  $x \in A$ , si ha  $[x]_\tau = \tilde{f}(\{f(x)\})$*

*Dimostrazione.* Sia  $x \in A$ . Allora  $[x]_\tau = \{y \in A \mid y \tau x\} = \{y \in A \mid f(y) = f(x)\}$ . Ora, per ogni  $y \in A$ , la condizione  $f(y) = f(x)$  è equivalente a  $f(y) \in \{f(x)\}$ , cioè alla condizione che  $y$  appartenga all'antiimmagine  $\tilde{f}(\{f(x)\})$  di  $\{f(x)\}$  mediante  $f$ . Pertanto  $[x]_\tau = \{y \in A \mid y \in \tilde{f}(\{f(x)\})\} = \tilde{f}(\{f(x)\})$ .  $\square$

**Teorema 12** (teorema di omomorfismo per insiemi). *Siano  $f : A \rightarrow B$  un'applicazione e  $\tau$  il suo nucleo di equivalenza. Allora, l'applicazione*

$$y \in \text{im } f \mapsto \tilde{f}(\{y\}) \in A/\tau$$

è biettiva ed ha per inversa l'applicazione

$$\tilde{f} : [x]_\tau \in A/\tau \mapsto f(x) \in \text{im } f.$$

Di conseguenza,  $|A/\tau| = |\text{im } f|$ .

*Dimostrazione.* Ricordiamo che  $\text{im } f$  è, per definizione, l'insieme  $\{f(x) \mid x \in A\}$ , quindi per ogni  $y \in \text{im } f$  esiste  $x \in A$  tale che  $y = f(x)$  e così, per il lemma 11,  $\tilde{f}(\{y\}) = [x]_\tau \in A/\tau$ . Questo mostra che l'applicazione  $\alpha : y \in \text{im } f \mapsto \tilde{f}(\{y\}) \in A/\tau$  è ben definita. Verifichiamo che essa è suriettiva. Per ogni  $c \in A/\tau$  esiste  $x \in A$  tale che  $c = [x]_\tau$ , e  $[x]_\tau = \tilde{f}(\{f(x)\})$  ancora per il lemma 11, inoltre  $f(x) \in \text{im } f$ , dunque  $c = \alpha(f(x))$ . Pertanto  $\alpha$  è suriettiva. Per verificare che  $\alpha$  è iniettiva siano ora  $u, v \in \text{im } f$  tali che  $\alpha(u) = \alpha(v)$ . Dal momento che  $u \in \text{im } f$ , esiste  $x \in A$  tale che  $u = f(x)$ . Per tale  $x$  si ha così  $x \in \tilde{f}(\{u\}) = \alpha(u)$ . Ma abbiamo assunto  $\alpha(u) = \alpha(v)$ , quindi si ha anche  $x \in \alpha(v) = \tilde{f}(\{v\})$ , cioè  $f(x) = v$ . Pertanto  $u = f(x) = v$ . Abbiamo così mostrato che  $\alpha$  è iniettiva, quindi anche biettiva. Vogliamo infine descrivere  $\alpha^{-1}$ . A questo scopo, sia  $c$  un qualsiasi elemento di  $A/\tau$ . Naturalmente  $c = [x]_\tau$  per un opportuno  $x \in A$ , e  $\alpha(f(x)) = \tilde{f}(\{f(x)\}) = [x]_\tau$  per il lemma 11, quindi  $f(x) = \alpha^{-1}([x]_\tau) = \alpha^{-1}(c)$ . In questo modo abbiamo verificato che l'inversa di  $\alpha$  è, come richiesto dall'enunciato, l'applicazione  $\tilde{f} : [x]_\tau \in A/\tau \mapsto f(x) \in \text{im } f$ , provando così anche che quest'applicazione è ben definita.  $\square$

La situazione descritta nel teorema precedente può essere descritta da questo diagramma

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ \pi_\tau \downarrow & & \uparrow \iota \\ A/\tau & \xrightarrow{\tilde{f}} & \text{im } f \end{array}$$

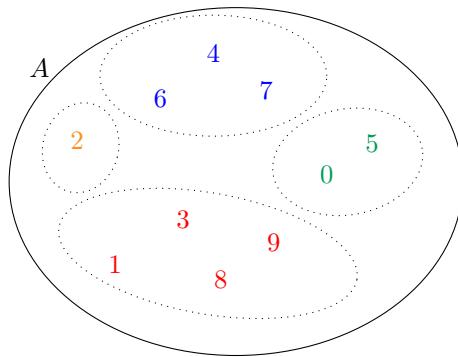
in cui  $\pi_\tau$  è la proiezione canonica definita dal nucleo di equivalenza  $\tau$  dell'applicazione  $f$ ,  $\tilde{f}$  è definita come nell'enunciato del teorema e  $\iota$  è l'immersione di  $\text{im } f$  in  $B$ , ed  $\tilde{f}$  è biettiva per il teorema 12. Risulta  $f = \iota \circ \tilde{f} \circ \pi_\tau$ , infatti per ogni  $x \in A$  si ha  $(\iota \circ \tilde{f} \circ \pi_\tau)(x) = \iota(\tilde{f}(\pi_\tau(x))) = \iota(\tilde{f}([x]_\tau)) = \iota(f(x)) = f(x)$ . Siccome  $\pi_\tau$  è suriettiva,  $\tilde{f}$  è biettiva e  $\iota$  è iniettiva, vediamo così che *ogni applicazione è ottenibile come composta tra un'applicazione iniettiva ed una suriettiva:  $f = (\iota \circ \tilde{f}) \circ \pi_\tau$ .*<sup>(5)</sup>

Vediamo qualche esempio:

- (1) Siano  $A = \{n \in \mathbb{Z} \mid -3 \leq n \leq 5\}$  e  $f: n \in A \mapsto n^2 \in \mathbb{N}$ . Sia poi  $\tau$  il nucleo di equivalenza di  $f$ . Per descrivere il quoziente  $A/\tau$  possiamo partire dalla descrizione di  $\text{im } f$ . Dovrebbe essere chiaro che  $\text{im } f = \{n^2 \mid n \in A\} = \{0, 1, 4, 9, 16, 25\}$ . Allora  $\text{im } f$  ha sei elementi, quindi, per il teorema 12 abbiamo  $|A/\tau| = 6$ ; in altri termini in  $A$  ci sono esattamente sei classi di equivalenza rispetto a  $\tau$ , che corrispondono ai sei elementi di  $\text{im } f$ . Le classi, cioè gli elementi di  $A/\tau$  sono dunque, sempre per lo stesso teorema:  $\tilde{f}(\{0\}) = \{0\}$ ,  $\tilde{f}(\{1\}) = \{1, -1\}$ ,  $\tilde{f}(\{4\}) = \{2, -2\}$ ,  $\tilde{f}(\{9\}) = \{3, -3\}$ ,  $\tilde{f}(\{16\}) = \{4\}$  e  $\tilde{f}(\{25\}) = \{5\}$ .
- (2) Utilizziamo lo stesso insieme  $A$  dell'esempio precedente, e studiamo il nucleo di equivalenza  $\sigma$  dell'applicazione  $g: x \in \mathcal{P}(A) \mapsto x \cap \mathbb{N} \in \mathcal{P}(\mathbb{N})$ . Iniziamo con l'identificare  $\text{im } g$ . Posto  $B = A \cap \mathbb{N}$ , per ogni  $x \in \mathcal{P}(B)$  si ha evidentemente  $g(x) = x \cap \mathbb{N} \subseteq B$ . Viceversa, per ogni  $y \in \mathcal{P}(B)$  abbiamo  $y = y \cap \mathbb{N} = g(y)$ , quindi  $y \in \text{im } g$ . Pertanto  $\text{im } g = \mathcal{P}(B)$  e quindi, come mostra il teorema 12,  $|\mathcal{P}(A)/\sigma| = |\text{im } g| = |\mathcal{P}(B)| = 2^{|B|} = 32$ . Come sono fatte le singole classi di equivalenza rispetto a  $\sigma$ ? Sempre per lo stesso teorema esse corrispondono agli elementi di  $\text{im } g = \mathcal{P}(B)$ , sono cioè gli insiemi  $\tilde{g}(\{y\})$  al variare di  $y \in \mathcal{P}(B)$ . Ad esempio, l'elemento  $\emptyset$  di  $\mathcal{P}(B)$  corrisponde alla classe  $\tilde{g}(\{\emptyset\}) = \{x \in \mathcal{P}(A) \mid g(x) = \emptyset\} = \{x \in \mathcal{P}(A) \mid x \cap \mathbb{N} = \emptyset\}$ . Non è difficile vedere che questo insieme è  $\mathcal{P}(A \setminus B) = \mathcal{P}(\{-3, -2, -1\})$ , un insieme di otto elementi. Chi legge può provare a dimostrare che, per ogni  $y \in \mathcal{P}(B)$  si ha  $\tilde{g}(\{y\}) = \{x \in \mathcal{P}(A) \mid g(x) = y\} = \{y \cup z \mid z \in \mathcal{P}(A \setminus B)\}$  e questo insieme ha esattamente otto elementi. Come ulteriore esempio, dal momento che  $g(\{1, -1\}) = \{1\}$ , il lemma 11 mostra che  $[\{1, -1\}]_\sigma = \tilde{g}(g(\{1, -1\})) = \tilde{g}(\{1\}) = \{\{1\} \cup z \mid z \in \mathcal{P}(A \setminus B)\}$ .

Per un esempio conclusivo, che illustri i diversi punti di vista presentati in queste note, facciamo ancora una volta riferimento allo stesso insieme  $A$ , attribuendo un colore ad i suoi elementi, come indicato qui:  $A = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ . Con riferimento ad  $A$  ed a questi colori possiamo ...:

- ... definire una relazione di equivalenza  $\rho \in \text{Eq}(A)$  dichiarando, per ogni  $x, y \in A$ ,  $x \rho y$  se e solo se  $x$  e  $y$  hanno lo stesso colore;
- ... definire un'applicazione  $c: A \rightarrow C$ , dove  $C$  è un insieme di colori contenente (almeno) i quattro colori utilizzati, ad esempio  $C = \{\text{verde, rosso, giallo, arancione, blu, grigio}\}$ . Come la nostra rappresentazione suggerisce,  $c$  manda 0 e 5 in verde, 1, 3, 8 e 9 in rosso, 4, 6 e 7 in blu e 2 in arancione.
- ... "raggruppare" gli elementi di  $A$  per colore, come nel diagramma che segue. Questo significa definire una partizione  $\mathcal{F}$  di  $A$  costituita da quattro blocchi, uno per ciascuno dei colori utilizzati, ciascuno dei quali è l'insieme degli elementi di  $A$  del colore dato, quindi  $\mathcal{F} = \{\{0, 5\}, \{1, 3, 8, 9\}, \{4, 6, 7\}, \{2\}\}$ :



Vediamo (e se non lo vediamo immediatamente riflettiamoci sopra sino a convincercene) che  $\rho$  non è altro che il nucleo di equivalenza di  $c$ , e che  $\mathcal{F}$  è precisamente  $A/\rho$ . Come sappiamo dal teorema 8,  $\mathcal{F}$  è determinata univocamente da  $\rho$  e  $\rho$  è determinata univocamente da  $\mathcal{F}$ , quindi assegnare  $\rho$  equivale ad assegnare  $\mathcal{F}$ . In accordo con il teorema 12,  $\mathcal{F} = A/\rho$  ha esattamente  $4 = |\text{im } c|$  elementi (le quattro classi rispetto a  $\rho$ , ovvero i quattro colori utilizzati). Notiamo infine che, mentre  $\mathcal{F}$  e  $\rho$  sono determinate da  $c$ ,  $c$  non è univocamente determinata da  $\mathcal{F}$  e  $\rho$ . Infatti, se proviamo a cambiare l'insieme  $C$  (il codominio di  $c$ ) aggiungendo ad esempio un nuovo colore (come viola), oppure cancellando da esso uno dei colori non usati (come grigio), l'applicazione  $c$  cambia, ma il suo nucleo

<sup>(5)</sup>o, se si preferisce,  $f = \iota \circ (\tilde{f} \circ \pi_\tau)$ .

di equivalenza resta  $\rho$ . Ancora più radicalmente, possiamo scambiare tra loro i colori, o sostituirne alcuni con altri in modo che l'applicazione  $c$  cambi senza che cambi il suo nucleo di equivalenza. Ad esempio, se ripetiamo l'intera discussione a partire da una colorazione di  $A$  come questa:  $A = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ , otteniamo un insieme  $C$  ed un'applicazione  $c$  sicuramente diversi da quelli presentati sopra ma, come si invita chi legge a verificare, la relazione d'equivalenza e la partizione risultanti (quelle che sopra erano  $\rho$  ed  $\mathcal{F}$ ) non cambieranno.

## Anelli

Un anello è una struttura algebrica  $(R, +, \cdot)$  tale che:

1.  $(R, +)$  sia un gruppo abeliano [il neutro si indica con  $0_R$ ]
2.  $(R, \cdot)$  sia un semigruppo
3.  $\cdot$  è distributivo rispetto a  $+$

$R$  è un anello commutativo  $\Leftrightarrow \cdot$  è commutativo.

$R$  è un anello unitario  $\Leftrightarrow (R, \cdot)$  ha el. neutro [ $1_R$ ].

$\forall T \subseteq S \quad (P(T), \Delta, \cap)$  sottoanello di  $(P(S), \Delta, \cap)$

Un sottoanello unitario deve conservare l'unità dell'anello ambiente.

$\mathbb{Z}$  sottoanello unitario di  $\mathbb{Q}$  con le op. indotte da  $\mathbb{R}$ .

$P(T)$  è un anello unitario ma non un sottoanello unitario di  $S$ .

Algebra

Lezione 14/11



## Anelli: tipi e caratteristiche

Siano  $R, S$  anelli unitari e sia  $f: R \rightarrow S$  un omomorfismo.

Non necessariamente  $f(1_R) = 1_S$ .

Esempio:

$$b \neq a \quad f: P(b) \xrightarrow{\text{iniezione}} P(a) \quad f(1_{P(b)}) = f(b) = b \neq a = 1_{P(a)}$$

Sia  $a$  un insieme. Se  $|a| \geq 1 \quad \exists X, Y \subseteq a \quad (X \neq \emptyset \neq Y \wedge X \cap Y = \emptyset)$   
equivalente a  $X, Y \in P(a)$

In  $P(a)$  NON vale la legge di annullamento del prodotto:

$$X \cap Y = \emptyset = O_{P(a)}, \text{ ma } X \neq O_{P(a)} \neq Y$$

Multiplicazione in  $P(a)$

$$\text{In } \mathbb{Z} \times \mathbb{Z}: (1, 0) \cdot (0, 1) = (0, 0) = O_{\mathbb{Z} \times \mathbb{Z}}$$

ma  $(1, 0) \neq O_{\mathbb{Z} \times \mathbb{Z}} \neq (0, 1)$

## Divisori dello zero

DEF Siano  $R$  un anello e  $a$  un suo elemento:

- $a$  è un divisore sx dello zero in  $R \iff \exists b \in R - \{O_R\} : (ab = O_R)$
- $a$  è un divisore dx dello zero in  $R \iff \exists b \in R - \{O_R\} : (ba = O_R)$
- $a$  è un divisore dello zero in  $R \iff$  è divisore sx o dx dello zero in  $R$ .

Se  $|R| \geq 1$   $O_R$  è un divisore dello zero (sia sx che dx).

Legge di annullamento del prodotto (o L.A.P.):

$$\forall a, b \in R - \{O_R\} : (ab \neq O_R) \iff R - \{O_R\} \text{ chiuso risp. a.}$$

L.A.P. vale se e solo se in  $R - \{O_R\}$  non esistono divisori dello zero.

Sia  $R$  un anello:

- $R$  è **integro**  $\Leftrightarrow \mathcal{O}_R$  è l'unico divisore dello zero in  $R$   $\Leftrightarrow |R| \geq 1$  e in  $R$  vale la L.A.P.
- $R$  è un **dominio di integrità**  $\Leftrightarrow R$  è integro e commutativo.

Esempi:  $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$  Non  $P(S)$  se  $|S| > 1$

Siano  $R$  un anello e  $a \in R$ . Allora:

- $a$  è un divisore sx dello zero in  $R$   $\Leftrightarrow a$  non è cancellabile a sx in  $R$ .
- $a$  è un divisore dx dello zero in  $R$   $\Leftrightarrow a$  non è cancellabile a dx in  $R$ .
- $a$  è un divisore dello zero in  $R$   $\Leftrightarrow a$  non è cancellabile in  $R$ .

DIM "=>" Sia  $a$  un divisore sx dello zero in  $R$ .

(caso sx) Allora  $\exists b \in R - \{\mathcal{O}_R\}$  ( $ab = \mathcal{O}_R$ ).

Fissato un tale  $b$  abbiamo:

$ab = \mathcal{O}_R = a \mathcal{O}_R$ , dunque  $ab = a \mathcal{O}_R$ , ma  $b \neq \mathcal{O}_R$

Quindi  $a$  non è cancellabile a sx in  $R$ .

"=<" Sia  $a$  non cancellabile a sx in  $R$ , cioè:  $\exists x, y \in R$  ( $ax = ay \wedge x \neq y$ )

Per tali  $x, y$   $ax - ay = \mathcal{O}_R \Rightarrow a(x-y) = \mathcal{O}_R$

Posto  $b = x-y$   $ab = \mathcal{O}_R \neq b$ , quindi  $a$  è divisore dello zero.

$R$  anello unitario e  $|R| \geq 1 \Rightarrow \mathcal{O}_R$  non è cancellabile  $\Rightarrow \mathcal{O}_R$  non è invertibile

•  $R$  è un **corpo**:  $\Leftrightarrow \begin{cases} R \text{ è un anello unitario.} \\ (R \text{ è integro}) \end{cases}$   $|R| \geq 1$  e ogni elemento di  $R - \{\mathcal{O}_R\}$  è invertibile (cioè  $\mathcal{U}(R) = R - \{\mathcal{O}_R\}$ )

•  $R$  è un **campo**:  $\Leftrightarrow R$  è un corpo commutativo.

(dominio di integrità) Esempi:  $\mathbb{R}, \mathbb{Q}, P(\{x\}) = \{\emptyset, \{x\}\} = \{\mathcal{O}_{P(\{x\})}, 1_{P(\{x\})}\}$

•  $R$  è **integro finito**  $\Rightarrow \begin{cases} R - \{\mathcal{O}_R\} \text{ chiuso rispetto a } \cdot; \\ (R - \{\mathcal{O}_R\}, \cdot) \text{ è un semigruppo finito} \end{cases}$

con tutti gli elementi sono cancellabili, quindi è un gruppo.

Allora  $R$  è un campo ( $\Rightarrow$  è un campo, in realtà).

## Binomio di Newton

Siano  $R$  un anello e  $a, b \in R$

$$(a+b)^2 = (a+b) \cdot (a+b) = a^2 + ab + ba + b^2$$

$$ab = ba \Rightarrow (a+b)^2 = a^2 + 2ab + b^2$$

$$(a+b)^3 = (a+b) \cdot (a+b) \cdot (a+b)$$

$ab = ba \Rightarrow (a+b)^3$  è una somma di prodotti nella forma  $a^i b^{3-i}$

$$a^3 = a \cdot a \cdot a$$

$$\leftarrow P_3(\{1, 2, 3\})$$

$$\binom{3}{3} = 1$$

$$a^2 b = a \cdot a \cdot b$$

$$a \cdot b \cdot a \quad b \cdot a \cdot a$$

$$\leftarrow P_2(\{1, 2, 3\})$$

$$\binom{3}{2} = 3$$

$$a b^2 = a \cdot b \cdot b$$

$$b \cdot a \cdot b \quad b \cdot b \cdot a$$

$$\leftarrow P_1(\{1, 2, 3\})$$

$$\binom{3}{1} = 3$$

$$b^3 = b \cdot b \cdot b$$

$$\leftarrow P_0(\{1, 2, 3\})$$

$$\binom{3}{0} = 1$$

$$(a+b)^3 = \sum_{i=0}^3 \binom{3}{i} a^i b^{3-i}$$

### Binomio di Newton

$\forall n \in \mathbb{N}^*$ , se  $ab = ba$  allora:

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

**Trivio:** Si chiama **coefficiente binomiale** perché è stato introdotto per svolgere operazioni come il quadrato e il cubo del **binomio**.

## Relazioni d'ordine

Sia  $A$  un insieme e  $p \in \text{Rel}(A)$   $p^\#$  = grafico di  $p$   $\bar{p}$  = rel. duale di  $p$ , con grafico  $(\bar{p})^\#$

- $p$  è **antiriflessiva** : $\Leftrightarrow \forall x \in A (x p x) \Leftrightarrow \Delta_A \cap p^\# = \emptyset$   
 $p$  non è antiriflessiva : $\Leftrightarrow \exists x \in A (x p x)$
- $p$  è **antisimmetrica** : $\Leftrightarrow \forall x, y \in A ((x p y \wedge y p x) \Rightarrow x = y) \Leftrightarrow p^\# \cap (\bar{p})^\# \subseteq \Delta_A$   
 $p$  non è antisimmetrica : $\Leftrightarrow \exists x, y \in A (x p y \wedge y p x \wedge x \neq y)$

•	•	
	•	

antisimmetrica  
non riflessiva  $\wedge$  non antiriflessiva

•		
	•	

Simmetrica  $\wedge$  antisimmetrica  
non riflessiva  $\wedge$  non antiriflessiva

	•	
		•

antisimmetrica  
 $\wedge$  antiriflessiva

- $p$  è una **relazione d'ordine largo** se e solo se è **riflessiva**, **antisimmetrica** e **transitiva**.

Esempi:  $\leq$  in  $\mathbb{R}$ ,  $\subseteq$  in qualche insieme

= **unica** relazione d'equivalenza e d'ordine

- $p$  è una **relazione d'ordine stretto** se e solo se è **antiriflessiva** e **transitiva**.

Esempi:  $<$  in  $\mathbb{R}$ ,  $\subset$  in qualche insieme

A A

$\forall p \in \text{OL}(A)$

ordine largo  $\hat{\wedge}$  definisco  $p_\# \in \text{Rel}(A)$ , ponendo:

$\forall x, y \in A, x p_\# y \Leftrightarrow (x p y \wedge x \neq y)$

Allora  $p_\# \in \text{OS}(A)$

ordine stretto

Infatti: 1)  $p_\#$  è **antiriflessiva**:  $\forall x \in A (x p_\# x)$

2)  $p_\#$  è **transitiva**:  $\forall x, y, z \in A$

$$(x p_\# y \wedge y p_\# z) \Rightarrow (x p y \wedge x \neq y) \wedge (y p z \wedge y \neq z) \Rightarrow (x p z \wedge x \neq z)$$

Altrimenti, se  $x = z$ , avremmo  $x p y \wedge y p x$ , dunque  $x = y$ .

# Algebra

## Lezione 16/11



## Relazioni d'ordine (all over again)

$\forall a \quad \forall p \in OL(a) \quad \forall \sigma \in OS(a)$

definiamo  $p_{\neq}, \sigma_{\equiv} \in Rel(a)$  ponendo:  $\forall x, y \in a$

$$x p_{\neq} y \iff (x p y \wedge x \neq y)$$

$$x \sigma_{\equiv} y \iff (x \sigma y \vee x = y)$$

DIM  $\forall \sigma \in OS(a)$

- $\sigma_{\equiv}$  è riflessiva: ovvio,  $\forall x \in a \quad x \sigma_{\equiv} x$

- $\sigma_{\equiv}$  è antisimmetrica:  $\forall x, y \in a \quad (x \sigma_{\equiv} y \wedge y \sigma_{\equiv} x \wedge x \neq y)$

$$\Rightarrow (x \sigma y \wedge y \sigma x \wedge x \neq y)$$

Assumo perché  $\sigma$  è antisimmetrica!

- $\sigma_{\equiv}$  è transitiva:  $\forall x, y, z \in a \quad (x \sigma_{\equiv} y \wedge y \sigma_{\equiv} z) \Rightarrow ((x \sigma y \vee x = y) \wedge (y \sigma z \vee y = z))$

Caso 1:  $x = y \vee y = z \Rightarrow x \sigma_{\equiv} z$

Caso 2:  $x \neq y \neq z \Rightarrow x \sigma y \wedge y \sigma z \Rightarrow x \sigma z$

per transitività



Le applicazioni:

$$p \in OL(a) \mapsto p_{\neq} \in OS(a)$$

$$\sigma \in OS(a) \mapsto \sigma_{\equiv} \in OL(a)$$

Sono biettive, una inversa dell'altra.

DIM

Per ogni  $p \in OL(a), \sigma \in OS(a)$

- se  $p^*$  è il grafico di  $p$ , quello di  $p_{\neq}$  è  $p^* \setminus \Delta_a$

- se  $\sigma^*$  è il grafico di  $\sigma$ , quello di  $\sigma_{\equiv}$  è  $\sigma^* \cup \Delta_a$

Allora: il grafico di  $(p_{\neq})_{\neq}$  è  $(p^* \setminus \Delta_a) \cup \Delta_a = p$

Similmente, il grafico di  $(\sigma_{\equiv})_{\neq}$  è  $(\sigma^* \cup \Delta_a) \setminus \Delta_a = \sigma$ .



## Divisibilità

Sia  $(S, \cdot)$  un semigruppo commutativo. La relazione di divisibilità in  $(S, \cdot)$ , indicata da  $|$  o da  $|_{(S, \cdot)}$ , è definita dalla formula:

$$\forall a, b \in S \quad (a|b \iff \exists c \in S (b = ac))$$

N.B. I concetti di divisione dello zero e divisibilità sono diversi!

### Esempi:

- $|_m (\mathbb{Z}, \cdot)$ :

$$3|6 \text{ sì} \quad 3|7 \text{ no} \quad -1|7 \text{ sì} \quad 3|0 \text{ sì}$$

- $|_m (2\mathbb{Z}, \cdot)$ :

$$2|6 \text{ no} \quad 2|8 \text{ sì}$$

- $|_m (P(\mathbb{Z}), \cup)$ :

$$\{\{2\}\} | \{\{6\}\} \text{ no} \quad \{\{2\}\} | \mathbb{N} \text{ sì}$$

- $(S, \cdot)$  è un semigruppo  $\Rightarrow |_{(S, \cdot)}$  è transitiva:

DIM  $\forall x, y, z \in S$

$$(x|y \wedge y|z) \Rightarrow \exists c, d \in S (y = xc \wedge z = yd) \\ \Rightarrow z = (xc)d = x(cd) \Rightarrow x|z$$

- $(S, \cdot)$  è un monide  $\Rightarrow |_{(S, \cdot)}$  è riflessiva:

DIM Se  $t$  è il neutro di  $(S, \cdot)$

$$\forall a \in S \quad a = at, \text{ quindi } a|a \quad \boxed{\text{sì}}$$

$|_{(\mathbb{Z}, \cdot)}$

$$1|-1 \quad -1|1 \quad \text{ma} \quad (-1)^{-1} \neq 1, \text{ quindi } |_{(\mathbb{Z}, \cdot)} \text{ non è antisimmetrica}$$

## Elementi associati

DEF  $\forall a, b \in S (a \sim b \Leftrightarrow ((a|b) \wedge (b|a)))$   
(Se  $a \sim b$  si dice che  $a$  e  $b$  sono **associati** in  $S$ )  $\sim_{(S, \cdot)}$

Se  $(S, \cdot)$  è un monoido,  $|_{(S, \cdot)}$  è d'ordine se e solo se è **antisimmetrica**, cioè se e solo se:  
 $\forall a, b \in S (a \sim b \Rightarrow a = b)$

N.B. L'implicazione  $a = b \Rightarrow a \sim b$  è ovvia perché  $|$  è riflessiva

Questo accade in  $(\mathbb{N}, \cdot)$ , quindi la **relazione di divisibilità** in  $(\mathbb{N}, \cdot)$  è d'ordine.



## Insieme ordinato

DEF È una coppia ordinata  $(S, p)$  dove  $p \in OL(S)$

### Esempi standard:

$(\mathbb{R}, \leq)$  ,  $(P(S), \subseteq)$  ,  $(\mathbb{N}, |)$   $\leftarrow$  **divisibilità  
in  $(\mathbb{N}, \cdot)$**

Siano  $a$  e  $b$  elementi di un insieme  $S$  in cui è definita una  $p \in OL(S)$ .

Diciamo che  $a$  e  $b$  sono **confrontabili** in  $(S, p)$  (oppure rispetto a  $p$ ) se esiste se  
 $a \leq b$  o  $b \leq a$  [negazione:  $a \not\leq b \wedge b \not\leq a$ ]

La relazione  $p$  è **TOTALE** se esiste se:

$\forall a, b \in S$   $a$  e  $b$  sono confrontabili rispetto a  $p$ ; in questo caso, diciamo che  $(S, p)$   
è **TOTALMENTE ORDINATO**. [negazione:  $\exists a, b \in S (a \not\leq b \wedge b \not\leq a)$ ]

## Esempi:

$(\mathbb{R}, \leq)$  è Tot. ordinato

$(P(S), \subseteq)$  è Tot. ordinato se e solo se  $|S| \leq 1$  (altrimenti basta considerare 2 elementi  
distinti  $x$  e  $y$  es  $\{x\} \neq \{y\}$  non confrontabili.)

$(\mathbb{N}, |)$  non è Tot. ordinato.  $2|3 \wedge 3|2$

Se  $S = \{0, 1\}$ ,  $(S, |_{(S, \cdot)})$  è Tot. ordinato.

## Insiemi d'ordine

Se  $(S, p)$  un insieme ordinato (quindi  $p \in OL(S)$ )

$\forall T \subseteq S$  si può definire la relazione binaria  $p_T$  indotta da  $p$  su  $T$ , via:

$\forall a, b \in T$  ( $a p_T b \Leftrightarrow a p b$ ) È ovvio che anche  $p_T$  è d'ordine:  $p_T \in OL(T)$

Allora  $(T, p_T)$  è un sottoinsieme ordinato di  $(S, p)$ .

In genere si scrive  $p$  anche per  $p_T$ , quindi facciamo riferimento al sottoinsieme ordinato  $(T, p)$  invece di  $(T, p_T)$

### Esempi:

$(\mathbb{R}^*, \leq), (\mathbb{Q}, \leq), (\mathbb{Z}, \leq), (\mathbb{N}, \leq)$  sono sottoinsiemi ordinati di  $(\mathbb{R}, \leq)$   
Se  $T \subseteq S$   $(P_{(T)}, \leq)$  lo è di  $(P_S, \leq)$

Se  $(S, p) \neq (T, \sigma)$  sono insiemi ordinati.

Un'applicazione  $f: S \rightarrow T$  è crescente da  $(S, p) \neq (T, \sigma)$  (o rispetto a  $p \neq \sigma$ )

se e solo se:  $\forall a, b \in S$   $a p b \Rightarrow f(a) \sigma f(b)$

Si dice che  $f$  è un isomorfismo da  $(S, p) \neq (T, \sigma)$  se e solo se è biiettiva, crescente da  $(S, p) \neq (T, \sigma)$  e, inoltre,  $f'$  è crescente da  $(T, \sigma) \neq (S, p)$ .

Questo equivale a:  $\forall a, b \in S$   $a p b \Leftrightarrow f(a) \sigma f(b)$

### Esempio:

L'applicazione identica su  $\mathbb{N}^*$ :  $\mathbb{N}^* \rightarrow \mathbb{N}^*$

è biiettiva e crescente da  $(\mathbb{N}^*, |)$  a  $(\mathbb{N}^*, \leq)$

ma non è crescente  $(\mathbb{N}^*, \leq)$  a  $(\mathbb{N}^*, |)$

$(\forall a, b \in \mathbb{N}^* \text{ } a | b \Rightarrow a \leq b)$

$(2 \leq 3 \text{ ma } 2 \nmid 3)$

# Algebra

## Lezione 18/11



# Massimi e minimi

Siano fissati  $S$  e  $p \in \text{OL}(S)$ . Sia  $\bar{p}$  la relazione duale a  $p$ .

$\forall x \in S$

$a$  è **minimo** in  $(S, p) \iff \forall x \in S (a \leq p x)$

$a$  è **minimale** in  $(S, p) \iff \neg (\exists x \in S (x \neq a \wedge p_x < p_a)) \iff \forall x \in S (p_a \Rightarrow x = a)$

$a$  è **massimo** in  $(S, p) \iff a$  è **minimo** in  $(S, \bar{p}) \iff \forall x \in S (p_a \geq x)$

$a$  è **massimale** in  $(S, p) \iff a$  è **minimale** in  $(S, \bar{p}) \iff \neg (\exists x \in S (a \leq p_x \wedge x \neq a)) \iff \forall x \in S (p_a \Rightarrow x = a)$

sì dice anche **rispetto a  $p$** , piuttosto che **in  $(S, p)$** .

N.B.

De non confondere con la nozione di (punto di) minimo / massimo di una funzione tra reali.

Se  $a$  è minimo in  $(S, p)$ , allora  $a$  è l'unico minimo in  $(S, p)$ .

DIM (1)  $a$  è minimo in  $(S, p)$ :

$\forall x \in S \ a \leq p x$  perché  $a$  è minimo.

Quindi, se  $x \neq a$ , allora  $x = a$  per antisimmetria.

(2) Se  $b$  è un elemento minimo in  $(S, p)$ , allora certamente  $b = a$ .

abbiamo  $a \leq b$  perché  $a$  è minimo.

$a \leq b \Rightarrow a = b$ , perché  $b$  è minimo, dunque  $b = a$ .



Per dualità: se  $a$  è massimo in  $(S, p)$ , allora  $a$  è l'unico elemento massimo in  $(S, p)$ .

Notazione:  $\min(S, p)$  e  $\max(S, p)$  indicano l'unico el. min e max. di  $(S, p)$  [potrebbero non esistere]

N.B.

Se  $(S, p)$  è tot. ordinato,  $\forall x \in S$   $a$  è minima  $\iff a$  è minimo;  $a$  è massima  $\iff a$  è massimo.

Esempi:

- $(\mathbb{N}, \leq)$  ha minimo 0, non ha minimale:  $\forall n \in \mathbb{N} (\exists x \in \mathbb{N} (n \leq x))$
- $(\mathbb{Z}, \leq)$  non ha minimale né massimali:  $\forall n \in \mathbb{Z} (\exists x \in \mathbb{Z} (\exists y \in \mathbb{Z} (x \leq n \wedge y \leq n)))$  ← non ha minimale ← non ha massimale
- $(P(S), \subseteq)$  minimo:  $\emptyset$  massimo:  $S$
- $(\mathbb{N}, |_{(\mathbb{N}, \cdot)})$  minimo: 1 massimo: 0 ( $\forall n \in \mathbb{N} ((1 \mid n \wedge n \mid 0))$ )
- $(\mathbb{N}, \#_{\mathbb{N}}, |_{(\mathbb{N}, \cdot)})$  minimale: i numeri primi (infiniti!) massimo: 0
- per ogni  $S$  in  $(S, \leq)$  ogni elemento è minima e massima ( $\forall a, b \in S (a \leq b \wedge b \leq a \iff a = b)$ )
- $p \in \text{OL}(\mathbb{Z})$ , definita da:  $\forall a, b \in \mathbb{Z} (a \leq p b \iff a \leq b \wedge (a = 0 \iff b = 0))$

In  $(\mathbb{Z}, p)$ , 0 è l'unico elemento minima e massima, ma non è min né max.

Se  $(S, p)$  insieme ordinato finito, ed  $a$  è l'unico minima, allora  $a = \min(S, p)$ .

## Teorema

Sia  $(S, \leq)$  un insieme ordinato finito non vuoto. Allora  $(S, \leq)$  ha elementi minimali ed elementi massimali.

DIM Supponiamo che  $(S, \leq)$  non abbia minimali.

Poiché  $S \neq \emptyset \exists x_0 \in S$ . Fissiamo  $x_0$ , non minima per assurto.

Allora esiste (e fissiamo)  $x_1 \in S : x_1 \nleq x_0$ .  $x_1$  non è minima

Allora esiste e fissiamo  $x_2 \in S : x_2 \nleq x_1$  e, per transitività  $x_2 \nleq x_0$ .

Allora  $\exists x_3 \in S : x_3 \nleq x_2 \wedge x_3 \nleq x_1 \wedge x_3 \nleq x_0$ .

Proseguendo in questo modo, definiamo una successione  $(x_n)_{n \in \mathbb{N}}$  di elementi di  $S$  tale che:

$\forall i, j \in \mathbb{N} (i < j \Rightarrow x_i \nleq x_j)$ , quindi  $\forall i, j \in \mathbb{N} (i \neq j \Rightarrow x_i \neq x_j)$

Allora  $\{x_i \mid i \in \mathbb{N}\}$  è una parte infinita di  $S$ , assurdo!!

Dunque  $S$  ha elementi minimali e, per dualità, massimali. (minimale rispetto a  $\bar{\leq}$ ). 



## Intervalli e copertura

Siano  $S$  un insieme,  $\leq$  e  $OL(S)$ ,  $<$  la corrispondente rel. d'ordine stretto,  $a, b \in S$ . Definiamo gli intervalli (chiuso, aperto, semicoperto):

$$[a, b] = \{x \in S \mid a \leq x \leq b\}$$

$$]a, b[ = \{x \in S \mid a < x < b\}$$

$$[a, b[ = \{x \in S \mid a \leq x < b\}$$

$$]a, b] = \{x \in S \mid a < x \leq b\}$$

Diciamo che  $b$  copre  $a$  (in  $(S, \leq)$ ) se e solo se  $a < b$  e  $]a, b[ = \emptyset$ .

In questo caso, scriviamo  $a < b$ .

Per ogni  $a, b \in S$ , sono equivalenti:

$$(1) a < b ;$$

$$(2) a < b \wedge \neg(\exists c \in S (a < c < b));$$

$$(3) a \text{ è un elemento massimale in } (\{x \in S \mid x < b\}, \leq);$$

$$(4) b \text{ è un elemento minima in } (\{x \in S \mid a < x\}, \leq).$$

## Esempi:

• In  $(\mathbb{Z}, \leq)$  2 è coperto da 3.  $\forall n \in \mathbb{Z}$ , n è coperto solo da  $n+1$ .

• In  $(\mathbb{Q}, \leq)$   $\forall a, b \in \mathbb{Q} (\neg(a < b))$ .

• In  $(P(\mathbb{N}), \subseteq)$ ,  $\emptyset$  è coperto da tutti e soli i singleton degli elementi di  $\mathbb{N}$ .

Se  $S$  è finito,  $\leq$  determina  $\leq$ , infatti:

### Teorema

Sia  $(S, \leq)$  un insieme ord. finito, e  $a, b \in S$ .

Allora  $a \leq b$  se esistono  $n \in \mathbb{N}$  elementi  $x_0, x_1, x_2, \dots, x_n$  di  $S$  tali che:

$$(1) \quad x_0 = a \quad (2) \quad x_n = b \quad (3) \quad \forall j \in \{i \in \mathbb{N} \mid i < n\} \quad (x_j \leq x_{j+1}).$$

### DIM

La condizione è sufficiente: se esistono  $n \in \mathbb{N}$  e gli elementi richiesti con le proprietà indicate, allora:  $n=0 \Rightarrow a=x_0=b$ , quindi  $a \leq b$ ;

$$n>0 \Rightarrow a=x_0 < x_1 < x_2 < x_3 \dots < x_n = b.$$

Quindi  $a \leq b$  in ogni caso.

Viceversa, se  $a \leq b$ , se  $a=b$ , allora ottieniamo la condizione richiesta ponendo  $n=0, x_0=a=b$ .

Se  $a \neq b$ , cioè  $a < b$ , allora poniamo  $x_0=a$  e consideriamo l'intervallo  $X = ]a, b]$ .

Se  $X$  è finito,  $X=\emptyset$  oppure  $(X, \leq)$  ha un elemento minima  $x_1$ .

Nel primo caso, se è  $a < b$ , quindi basta porre  $n=1$  e  $x_1=b$ .

Nel secondo caso, consideriamo l'intervallo  $]x_1, b]$  ( $x_1 \leq b$ ).

Se  $]x_1, b] = \emptyset$ , allora  $a = x_0 < x_1 < x_2 = b$  e la cond. è soddisfatta.

Se  $]x_1, b] \neq \emptyset$ , allora c'è  $x_2$  elemento minima rispetto a  $\leq$ ; allora  $a = x_0 < x_1 < x_2 \leq b$  e consideriamo  $]x_2, b]$ .

Procediamo ricorsivamente in questo modo, ottenendo una successione  $a = x_0 < x_1 < x_2 \dots < x_i < x_{i+1} \dots \leq b$ .

Poiché  $S$  è finito, ad un certo punto la successione dovrà fermarsi; quindi troviamo un  $i \in \mathbb{N}$ :  $x_i = b$ .

Ponendo  $n=i$ , abbiamo costruito gli elementi  $x_0, x_1, \dots, x_n$  richiesti.



Gli insiemi ordinati finiti sono codificati dalla loro relazione di copertura.

# Diagramma di Hasse

$(S, \leq)$  ins. ordinato finito.

Rappresentiamo gli elementi di  $S$  come punti del piano, col vincolo che se  $a, b \in S$  e  $a \leq b$ , il punto che rappresenta  $b$  sia disegnato più in alto di quello che rappresenta  $a$ . Inoltre tracciamo le linee da  $a$  a  $b$  se e solo se  $a < b$ .

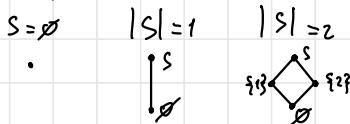
La figura così ottenuta è un diagramma di Hasse di  $(S, \leq)$ .

## Esempi:

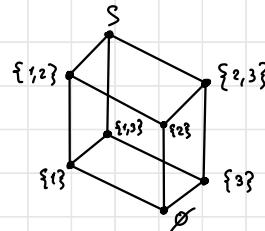
- $S = \{1, 2, 3\}$   $\leq$ : ordine usuale



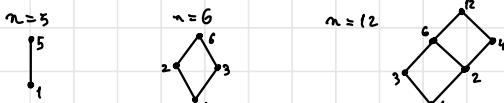
- $(P(S), \subseteq)$



$$|S|=3$$



- $\text{Div}(n) = \{d \in \mathbb{N} \mid d \mid n\}$  (è l'intervallo  $[1, n]$  in  $(\mathbb{N}, |)$ )



$$(\text{Div}(5), |) \cong (P(\{5\}), \subseteq)$$

$\Downarrow$   
isomorfismo

$$P(\{1, 2\}, \subseteq) \cong (\text{Div}(6), |)$$

# Algebra

## Lezione 21/11



## Minimali e Massimali (pt. 2)

$$f: S \rightarrow T \quad \text{app.} \quad \leq \in \text{OL}(T) \quad \leq = (\leq)_{\neq}$$

Definiamo  $\sigma, p \in \text{Rel}(S)$ , usando le formule:

$$\begin{array}{ll} \forall x, y \in S & x \sigma y \Leftrightarrow f(x) < f(y) \\ (p = \sigma_{\leq}) & x p y \Leftrightarrow (x = y \vee \underbrace{f(x) < f(y)}_{x \sigma y}) \end{array}$$

$$\forall x, y, z \in S$$

antiriflessiva:  $f(x) \not\sim f(x)$

transitiva:  $(\begin{matrix} x \sigma y \\ y \sigma z \end{matrix}) \Rightarrow (\begin{matrix} f(x) < f(y) \\ f(y) < f(z) \end{matrix}) \Rightarrow f(x) < f(z) \Leftrightarrow x \sigma z$

Allora  $\sigma \in \text{OS}(S)$  e  $p \in \text{OL}(S)$ .

$\forall a \in S \quad a \text{ è minimaile in } (S, p) \Leftrightarrow f(a) \text{ è minimaile in } (\text{inf}, \leq).$

$a$  non è minimaile in  $(S, p) \Leftrightarrow \exists b \in S (b p_{\neq} a) \Leftrightarrow \exists b \in S (b \sigma a)$

$\Leftrightarrow \exists b \in S (b \sigma a) \Leftrightarrow \exists b \in S (f(b) < f(a)) \Leftrightarrow f(a) \text{ non è minimaile in } (\text{inf}, \leq).$

Esempio:

$$f: x \in P(A) \mapsto |x| \in \mathbb{N} \quad (\mathbb{N}, \leq)$$

$$|A|=4 \quad \text{inf} = \left\{ \begin{matrix} \text{min} & 0, 1, 2, 3, 4 \end{matrix} \right\} \quad \text{max, NON sarebbe il massimo in } \mathbb{N}$$



Operazione duale con i diagrammi di Hasse



## Minorante e Maggiorante

$$(S, p) \quad p \in \mathcal{O}L(S)$$

$$\forall X \in P(S)$$

$$X_{(S,p)}^{\downarrow} = \left\{ a \in S \mid \forall x \in X \ (a \leq_p x) \right\}$$

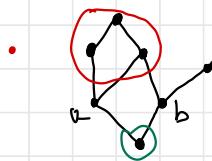
$$X_{(S,p)}^{\uparrow} = \left\{ a \in S \mid \forall x \in X \ (x \leq_p a) \right\}$$

$$\forall a \in S$$

$a$  è un minorante di  $X$  in  $(S, p)$   
se esiste  $a \in X_{(S,p)}^{\downarrow}$

$a$  è un maggiorante di  $X$  in  $(S, p)$   
se esiste  $a \in X_{(S,p)}^{\uparrow}$

Esempi:



$$\begin{matrix} \{a, b\}^{\uparrow} \\ \{a, b\}^{\downarrow} \end{matrix}$$

$$\bullet \quad l_n (\mathbb{Z}, \leq)$$

$$\mathbb{N}^{\uparrow} = \emptyset$$

$$\mathbb{N}^{\downarrow} = \{n \in \mathbb{Z} \mid n \leq 0\}$$

$$\bullet \quad l_n (P(S), \subseteq) \quad \forall S$$

$$\forall X \in P(S)$$

$$X^{\uparrow} = \left\{ Y \mid \forall a \in X \ (a \subseteq Y) \right\} =$$

$$= \left\{ Y \in P(S) \mid \forall a \in X \ (a \subseteq Y) \right\} =$$

$$\rightarrow \text{Esempio particolare} \quad S = \mathbb{N} \quad X = \left\{ \underbrace{\{1, 2\}}_A, \underbrace{\{2, 3\}}_B \right\}$$

dal particolare  
al generale

$$\begin{aligned} X^{\uparrow} &= \left\{ Y \in P(\mathbb{N}) \mid A \subseteq Y \wedge B \subseteq Y \right\} \\ &= \left\{ Y \in P(\mathbb{N}) \mid \forall c \in A \cup B \ (c \in Y) \right\} \\ &= \left\{ Y \in P(\mathbb{N}) \mid A \cup B \subseteq Y \right\} \end{aligned}$$

$$X^{\uparrow} = \left\{ Y \subseteq P(S) \mid \forall c \in UX \ (c \in Y) \right\} = \left\{ Y \subseteq S \mid UX \subseteq Y \right\}$$

$$X^{\downarrow} = \left\{ Y \subseteq P(S) \mid \forall a \in X \ (Y \subseteq a) \right\} =$$

$$\begin{aligned} \text{caso 1: } X^{\downarrow} &= \left\{ Y \subseteq P(S) \mid Y \subseteq \cap X \right\} \\ (X \neq \emptyset) \quad &= P(\cap X) \end{aligned}$$

$$\text{caso 2: } X^{\downarrow} = \left\{ Y \mid Y \subseteq P(S) \right\} = P(S)$$

$$\bullet \quad \emptyset^{\uparrow(S,p)} = S = \emptyset^{\downarrow(S,p)}$$

## Estremo inferiore ed Estremo superiore

$\forall a \in S$

$a$  è estremo inferiore di  $X$  in  $(S, p)$   $\Leftrightarrow a = \max_{(S, p)}$

$a$  è estremo superiore di  $X$  in  $(S, p)$   $\Leftrightarrow a = \min_{(S, p)}$

$X^{\downarrow(S, p)}$

$\Leftrightarrow a = \inf_{(S, p)}(X)$

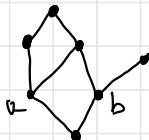
$X^{\uparrow(S, p)}$   $\Leftrightarrow a = \sup_{(S, p)}(X)$

$\Leftrightarrow a = \inf_{(S, p)} X^{\downarrow(S, p)}$

## Reticolo

Sia  $(S, p)$  un insieme ordinato.

Si dice RETICOLO se e solo se  $\forall a, b \in S$  esistono  $\inf_{(S, p)}(\{a, b\})$  e  $\sup_{(S, p)}(\{a, b\})$



Non è un reticolo, non ha sup ( $\{a, b\}$ )

# Algebra

## Lezione 23/11



# Reticoli

**Teorema.** Sia  $(S, p)$  un reticolo, sia  $a \in S$ .  
 $a$  minimale in  $(S, p) \Leftrightarrow a = \inf_{(S, p)} (S, p)$

DIM Sia  $a$  minimale.

$\forall b \in S$ , sia  $c = \inf \{a, b\}$ , allora

$$c \leq a \text{ e } c \leq b \Rightarrow c = a. \text{ Dunque } a = \inf_{(S, p)} (S, p).$$



Il duale di un reticolo è anch'esso un reticolo, con sup e inf scambiati rispetto a  $p$ .



$\forall a, b \in S$

$$a \leq b \Leftrightarrow a = \inf_{(S, p)} (a, b) \Leftrightarrow b = \sup_{(S, p)} (a, b)$$

Un reticolo si dice **completo** se ogni suo sottoinsieme ha inf e sup.

Un reticolo si dice **limitato** se ha min e max.

Un reticolo **completo** è anche **limitato**, ma non è vero il viceversa.

Sia  $X \subseteq S$  e  $x = \inf_{(S, p)} X$

$$\text{allora } X^{\downarrow} = \{x\}^{\downarrow}$$

$(S, p)$  ins. ordinato

$X, Y \subseteq S$

supponiamo esistano  $x = \inf_{(S, p)} (X)$  e  $y = \sup_{(S, p)} (Y)$

Allora:

$$(X \cup Y)^{\downarrow} = X^{\downarrow} \cap Y^{\downarrow} = \{x\}^{\downarrow} \cap \{y\}^{\downarrow} = \{x, y\}^{\downarrow}$$

Se  $(S, p)$  è un reticolo, allora:

$$\text{esiste } a = \inf \{x, y\} = \max (\{x, y\}^{\downarrow}) = \max ((X \cup Y)^{\downarrow}) = \inf (X \cup Y)$$

*Dimostrazione bonus per gli interessati*

$$\begin{aligned} (X \cup Y)^\downarrow : X^\downarrow \cap Y^\downarrow &= \left\{ \alpha \mid \forall z (z \in X \Rightarrow \alpha_p z) \right\} \cap \left\{ \alpha \mid \forall z (z \in Y \Rightarrow \alpha_p z) \right\} \\ \left\{ \alpha \mid \forall z \in X \cup Y \quad \alpha_p z \right\} &\qquad\qquad\qquad \text{..} \\ \left\{ \alpha \mid \forall z \left( (z \in X \vee z \in Y) \Rightarrow \alpha_p z \right) \right\} &\qquad\qquad\qquad \left\{ \alpha \mid \forall z (z \in X \Rightarrow \alpha_p z \wedge z \in Y \Rightarrow \alpha_p z) \right\} \\ \left\{ \alpha \mid \forall z (\neg(z \in X \vee z \in Y) \vee \alpha_p z) \right\} &\qquad\qquad\qquad \text{..} \\ \left\{ \alpha \mid \forall z (z \notin X \wedge z \notin Y) \vee \alpha_p z \right\} & \end{aligned}$$

## Esempio:

$(S, \leq)$  reticolo

$a, b, c \in S$

$$\exists i = \inf\{a, b\}$$

$$\text{oppure: } \exists j = \inf\{b, c\}$$

$$\exists c = \inf\{c\}$$

$$\inf\{a, j\} = \inf\{a, b, c\}$$

$$\exists \inf\{a, b, c\} = \inf\{i, c\}$$

Vale per tutti gli insiemi finiti non vuoti.

Se un insieme finito è un reticolo, allora è un reticolo completo.



## Meet e Join

$(S, \leq)$  reticolo.

Definiamo due operazioni binarie in  $S$ :

$\cap$ , wedge, meet $\rightarrow$	$\wedge : (a, b) \in S \times S \mapsto \inf_{(\leq)}\{a, b\} \in S$	intersezione reticolare
$\cup$ , vee, join $\rightarrow$	$\vee : (a, b) \in S \times S \mapsto \sup_{(\leq)}\{a, b\} \in S$	unione reticolare

(1)  $\vee$  e  $\wedge$  sono commutative:  $a \wedge b = \inf\{a, b\} = \inf\{b, a\} = b \wedge a$

(2)  $\vee$  e  $\wedge$  sono associative:  $(a \wedge b) \wedge c = a \wedge c = \inf\{\inf\{a, b\}, c\} = a \wedge (b \wedge c)$

(3) Valgono le leggi di assorbimento:  $\forall a, b \in S \quad a \wedge (a \vee b) = a = a \vee (a \wedge b)$

(4)  $\vee$  e  $\wedge$  sono idempotenti: ogni elemento è idempotente  $\forall a \in S \quad a \wedge a = a = a \vee a$

$\forall a \in S \quad a$  è neutro risp  $\vee \iff \forall b \in S \quad (b = a \vee b) \iff \forall b \in S \quad (a \leq b) \iff a = \min(S, \leq)$

Vale anche il duale  $a$  è neutro risp.  $\wedge \iff a = \max(S, \leq)$

Viceversa, sia  $(S, \vee, \wedge)$  una struttura algebrica con due op. binarie  $\vee$  e  $\wedge$  che verificano (1), (2), (3)

Definiamo  $p \in \text{Rel}(S)$  con la formula:  $\forall a, b \in S \quad a \vee p b \iff a = a \vee b$

per (1) + (3)

Osserviamo  $\forall a, b \in S \quad (a \vee p b \iff b = a \vee b)$ . Infatti, se  $a \vee p b$  allora  $a = a \vee b$ , quindi  $a \vee b = (a \vee b) \vee b = b$ .

$\forall a, b, c \in S \quad$  rifl.  $a \vee a$ , infatti  $(1, 2, 3) \Rightarrow (4)$ , quindi  $a = a \vee a$

antisim.  $(a \vee b \neq b \vee a) \Rightarrow a = a \vee b = b \vee a = b$

Assoc.  
(2)

trans.  $(a \vee b \neq b \vee c) \Rightarrow (a = a \vee b \neq b \vee c) \Rightarrow a = a \vee (b \vee c) = (a \vee b) \vee c = a \vee c \iff a \vee c$

$\forall a, b, c \in S$

$$M = \{a, b\}^{\downarrow(S, P)}$$

$$c \in M \Leftrightarrow (c \text{ p } a \wedge c \text{ p } b) \Leftrightarrow c = c \wedge a = c \wedge b \Rightarrow c \wedge (a \wedge b) = (c \wedge a) \wedge b = c \wedge b = c \Rightarrow c_p(a \wedge b)$$

$$\begin{aligned} (a \wedge b) \wedge a &= a \wedge (a \wedge b) = (a \wedge a) \wedge b = a \wedge b \\ (a \wedge b) \wedge b &= a \wedge (b \wedge b) = a \wedge b \end{aligned} \quad \left. \begin{array}{l} \\ \end{array} \right\} a \wedge b \in M$$

$$a \wedge b = \max_{(S, P)} M = \inf_{(S, P)} \{a, b\} \quad ; \text{ in modo analogo: } a \vee b = \sup_{(S, P)} \{a, b\}$$

Allora  $(S, P)$  è un reticolo con operazioni reticolari  $\wedge$  (per  $\inf$ ) e  $\vee$  (per  $\sup$ ).

Esempio:

$$X, Y \in P(S) \quad (P(S), \subseteq)$$

$$\begin{aligned} a \text{ p } b &\Leftrightarrow a = a \wedge b & X = X \wedge Y &\Leftrightarrow X \subseteq Y \\ \neg a \text{ p } b &\Leftrightarrow b = a \vee b & Y = X \vee Y &\Leftrightarrow X \supseteq Y \end{aligned}$$

## Isomorfismo tra reticolati

$$(P(\{1,2\}), \subseteq) \text{ ovvero } (P(S), \cup, \cap)$$



$$(\text{Div}(S), |)$$

Siano  $(S, \leq)$  e  $(T, \tau)$  reticolati con op. reticolari  $\vee, \wedge$  per  $S$  e  $\dot{\vee}, \dot{\wedge}$  per  $T$ .

Sia  $f: S \rightarrow T$  un'app. bisettiva.

Allora  $f$  è un isomorfismo di insiemi ordinati  $(S, \leq) \rightarrow (T, \tau)$

Se e solo se è un isomorfismo di strutture algebriche  $(S, \vee, \wedge) \rightarrow (T, \dot{\vee}, \dot{\wedge})$

$f$  isomorfismo di insiemi ordinati  $\Rightarrow \forall a, b \in S \quad a \wedge b = \inf_{(S, \leq)} \{a, b\}$

Ha per immagine

$$\begin{aligned} f(a \wedge b) &= \inf_{(T, \tau)} \{f(a), f(b)\} = f(a) \dot{\wedge} f(b) \\ f(a \vee b) &= f(a) \dot{\vee} f(b) \end{aligned} \Rightarrow$$

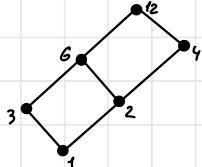
$f$  isomorfismo di strutture algebriche

$f$  isomorfismo di strutture algebriche  $\Rightarrow \forall a, b \in S \quad f(a \wedge b) = f(a) \dot{\wedge} f(b)$

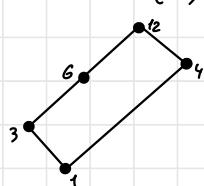
$$a \leq b \Leftrightarrow a = a \wedge b \Leftrightarrow f(a) = f(a) \dot{\wedge} f(b) \Leftrightarrow f(a) \tau f(b) \Rightarrow$$

$f$  isomorfismo di insiemi ordinati

$$S = \text{Div}(12)$$



$$T = S - \{2\}$$



è un reticolo, ma non un sottoreticolo di  $(S, |)$ .

Non è chiuso rispetto a  $|$  ( $\inf$  in  $S$ ).

$$6 \wedge 4 = 2 \notin T !!$$

$$\forall a, b \in S \quad a \leq b$$

$I = [a, b]$  è un sottoreticolo

$$\forall x, y \in I \quad a \leq x \wedge y \leq x \vee y \leq b$$

$$\in I$$

$$T \subseteq S \quad P(T) = [\emptyset, T]_{(P(S), \subseteq)}$$

## Complemento

$(S, \leq)$  reticolo limitato

a è un complemento di b in  $(S, \leq)$  (e quindi b è complemento di a)  
se e solo se  $a \wedge b = \min(S, \leq)$  e  $a \vee b = \max(S, \leq)$

Esempio:

$\min S$  e  $\max S$  sono l'uno complemento dell'altro.

~ Unico caso in cui a e b possono essere complementi tra loro e confrontabili. ~

Un reticolo si dice **complementato** se e solo se ogni elemento ha complemento.

$(P(S), \subseteq)$   $\forall X \in P(S) \quad S \setminus X \text{ è compl. di } X.$

# Algebra

## Lezione 25/11



## Distributività nei reticolati

Sia  $(S, \leq)$ , con oper. reticolari:  $\bigvee_{(\text{sup})}$  e  $\bigwedge_{(\text{inf})}$

$\forall a, b, c \in S$

$(S, \leq)$  è distributiva:  $\Leftrightarrow$   $\vee$  è distr. rispetto a  $\wedge$  e viceversa  $\Leftrightarrow a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$   
 $a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$

### Esempi:

- $(P(S), \subseteq)$
- Tutti i sottoreticolati di un reticolo distributivo
- Tutti i reticolati cui sottoreticolati sono tutti distributivi
- $(\mathbb{N}, |)$
- Tutti i reticolati completi.

Se  $(S, \leq)$  è distributivo (e quindi limitato) e  $a \in S$ , allora  $a$  ha al più un complemento in  $(S, \leq)$ .

DIM Siano  $x, y$  comp. di  $a$  in  $(S, \leq)$ . Sia  $0 = \min(S, \leq)$  e  $1 = \max(S, \leq)$

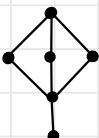
Allora  $a \wedge x = 0 = a \wedge y$ ,  $a \vee x = 1 = a \vee y$

$$x = x \wedge 1 = x \wedge (a \vee y) = (x \wedge a) \vee (x \wedge y) = 0 \vee (x \wedge y) = x \wedge y$$

$$\text{Similmente: } y = y \wedge x = x$$

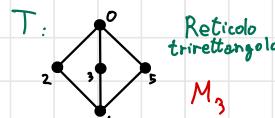
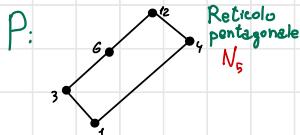
**N.B.** Esistono reticolati non distributivi i cui elementi, se hanno complemento, ne hanno uno solo.

### Esempio:

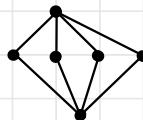
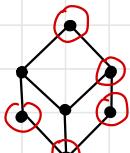
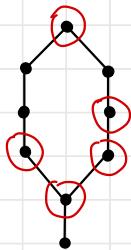


## Teorema di Birkhoff (criterio)

Sia  $(S, \leq)$  un reticolo. Allora  $S$  è distributivo se e solo se nessun sottoreticolo di  $(S, \leq)$  è isomorfo ad uno tra  $P$  e  $T$ .



Esempio:



Basta togliere  
uno dei punti: 4

# STRUTTURE BOOLEANE

GIOVANNI CUTOLO

Lo scopo di queste note è quello di presentare in modo unitario anelli booleani, reticolari booleani e algebre di Boole, senza entrare in troppi dettagli ma spiegando come e perché lo studio di ciascuna di queste strutture è equivalente a quello delle altre. Il riquadro che segue contiene un riassunto di questi contenuti; sia bene inteso che questo riassunto non è di per sé sufficiente per la loro comprensione, ma la sua lettura è un utile preliminare a quella del resto di queste note. Nella sezione finale delle note faremo poi qualche osservazione su come si possano inquadrare in questa teoria esempi di algebre di Boole che chi legge ha probabilmente incontrato, o sta per incontrare, in altri corsi.

## In sintesi

Si definiscono tre tipi di strutture che fanno riferimento nel loro nome a quello di George Boole. Abbiamo gli *anelli booleani*, che sono per definizione gli anelli unitari i cui elementi sono tutti idempotenti, i *reticolari booleani*, che sono invece i reticolari **distributivi** e **complementati**, le *algebre di Boole*, che sono particolari strutture algebriche la cui definizione è riportata [più avanti](#), nella terza sezione di queste note.

Ciò che lega queste strutture tra loro è che definire su un insieme una struttura di uno di questi tre tipi (anello booleano, reticolo booleano, algebra di Boole) equivale definirne una di ciascuno degli altri due tipi; in modo che risulti del tutto equivalente lo studio degli anelli booleani, quello delle algebre di Boole e quello dei reticolari booleani.

L'esempio da avere come riferimento è quello dell'insieme  $\mathcal{P}(S)$  delle parti di un insieme  $S$ . Come dovrebbe essere ben noto,  $(\mathcal{P}(S), \subseteq)$ , cioè l'insieme  $\mathcal{P}(S)$  ordinato per inclusione, è un reticolo, che risulta essere un reticolo booleano. Lo stesso insieme, munito delle operazioni di differenza simmetrica ed intersezione,  $(\mathcal{P}(S), \Delta, \cap)$ , è un anello booleano. Infine,  $\mathcal{P}(S)$  si può strutturare come algebra di Boole mediante le operazioni di unione, intersezione e l'operazione unaria di complemento  ${}^c$ :  $X \in \mathcal{P}(S) \mapsto S \setminus X \in \mathcal{P}(S)$ ; l'algebra di Boole così ottenuta è  $(\mathcal{P}(S), \cup, \cap, \emptyset, S, {}^c)$ .

Questo esempio è particolarmente importante per almeno due motivi. Uno di tipo pratico: il modo in cui si può, in  $\mathcal{P}(S)$ , passare da uno dei tre tipi di struttura booleana a ciascuno degli altri due illustra molto bene come si può effettuare l'analogico passaggio a partire da una struttura booleana arbitraria; questo esempio può essere quindi di grande aiuto nello studio della situazione generale. Il secondo motivo, di carattere teorico e di importanza ancora maggiore, è che quello fornito dagli insiemi  $\mathcal{P}(S)$  non è un esempio particolare ma, in qualche modo, quello tipico. Infatti un importante teorema (dovuto a M.H. Stone) mostra che ogni anello booleano finito è isomorfo a  $(\mathcal{P}(S), \Delta, \cap)$  per un opportuno insieme  $S$  (per gli anelli infiniti il teorema è un po' più debole: ogni anello booleano è isomorfo ad un sottoanello unitario di  $(\mathcal{P}(S), \Delta, \cap)$ , per un opportuno insieme  $S$ ). Analoghi enunciati valgono per i reticolari booleani e per le algebre di Boole. Questo vuol dire, ad esempio, che se sappiamo descrivere il reticolo delle parti degli insiemi finiti, conosciamo, a meno di isomorfismi, tutti i reticolari booleani finiti. Una conseguenza del teorema di Stone è che gli anelli booleani finiti (ma lo stesso vale per i reticolari booleani finiti o per le algebre di Boole finite) hanno per cardinalità una potenza di 2, e che due anelli booleani finiti con lo stesso numero di elementi sono necessariamente isomorfi.

*Avvertenza.* Alcune parti di questo file, in cui appaiono di regola dimostrazioni o verifiche, sono indentate e marcate da un segnale di pericolo. Questo indica che i loro contenuti vanno considerati approfondimenti per chi fosse ad essi interessato ma non fanno parte del programma del corso e non sono richiesti ai fini dell'esame. Altre osservazioni e dimostrazioni possono essere o non essere parte effettiva del programma, a seconda che siano o non siano state trattate a lezione.

## 1. ANELLI BOOLEANI

Per definizione un *anello booleano* è un anello unitario in cui ogni elemento è *idempotente*, cioè coincide col proprio quadrato.

Ad esempio, l'anello  $\mathbb{Z}_2$  degli interi modulo 2 è un anello booleano: è unitario e i suoi due elementi,  $\bar{0} = [0]_2$  e  $\bar{1} = [1]_2$  sono idempotenti:  $\bar{0}^2 = \bar{0}$  e  $\bar{1}^2 = \bar{1}$ . Un altro esempio significativo è quello dell'anello  $(\mathcal{P}(S), \Delta, \cap)$  delle parti di un (arbitrario) insieme  $S$ . Infatti quest'anello è unitario (di unità  $S$ ) e, dal momento che l'operazione di moltiplicazione nell'anello  $\mathcal{P}(S)$  è quella di intersezione, per ogni  $X \in \mathcal{P}(S)$  si ha  $X^2 = X \cap X = X$ .

Prima di dimostrare una semplice proprietà degli anelli booleani è opportuno un richiamo sulla nozione di caratteristica di un anello unitario. Se  $R$  è un anello unitario e l'unità  $1_R$  di  $R$ , ha periodo finito  $c$  nel gruppo additivo  $(R, +)$ , si dice che  $c$  è la *caratteristica* di  $R$ . Detto in modo più esplicito, se esiste qualche intero positivo  $n$  tale che  $n1_R$  (che è la somma  $1_R + 1_R + \dots + 1_R$  con  $n$  addendi) è uguale a  $0_R$  (lo zero di  $R$ ), allora la caratteristica

di  $R$  è il minimo tale intero  $n$ .<sup>(1)</sup> Dovrebbe essere chiaro che  $R$  ha caratteristica 1 se e solo  $1_R = 0_R$ ; si verifica facilmente che in questo caso  $R = \{0_R\}$ . Il caso immediatamente successivo è quello degli anelli di caratteristica 2: sono quelli in cui  $1_R \neq 0_R$  ma  $2 \cdot 1_R = 1_R + 1_R = 0_R$ . Notiamo che l'anello  $(\mathcal{P}(S), \Delta, \cap)$  ha questa proprietà se  $S \neq \emptyset$ . Infatti in questo anello l'unità è  $S$ , lo zero è  $\emptyset$ , l'addizione è l'operazione di differenza simmetrica e si ha  $2 \cdot S = S \Delta S = \emptyset$ . Quindi l'anello  $\mathcal{P}(S)$  ha caratteristica 2.

Dimostriamo ora che quanto appena visto per  $(\mathcal{P}(S), \Delta, \cap)$  vale per tutti gli anelli booleani; verificando inoltre che questi anelli sono sempre commutativi.

**Proposizione 1.** *Sia  $R$  un anello booleano. Allora  $R$  è commutativo e, se  $|R| > 1$ ,  $R$  ha caratteristica 2.*

*Dimostrazione.* Per ogni  $a, b \in R$  si ha  $a^2 = a$ ,  $b^2 = b$  e  $(a + b)^2 = a + b$ , perché  $R$  è booleano. D'altra parte, come in ogni anello,

$$(a + b)^2 = (a + b)(a + b) = a(a + b) + b(a + b) = a^2 + ab + ba + b^2$$

e quindi

$$a + b = (a + b)^2 = a^2 + ab + ba + b^2 = a + ab + ba + b.$$

Da ciò, cancellando  $a$  e  $b$ , si ricava  $ab + ba = 0_R$ . Dunque:

$$(\forall a, b \in R) (ab = -ba). \quad (*)$$

Applicando la  $(*)$  nel caso in cui  $a = b$  si ottiene, per ogni  $a \in R$ ,  $a^2 = -a^2$ . Ma  $a^2 = a$ , quindi:

$$(\forall a \in R) (a = -a); \quad \text{ovvero:} \quad (\forall a \in R) (2a = 0_R). \quad (**)$$

In particolare,  $2 \cdot 1_R = 0_R$ , quindi o  $1_R = 0_R$  e  $R = \{0_R\}$  ha un solo elemento, oppure  $|R| > 1$  e la caratteristica di  $R$  è 2.

Infine, per ogni  $a, b \in R$ , applicando la  $(**)$  all'elemento  $ba$  otteniamo  $-ba = ba$ , quindi la  $(*)$  prova  $ab = ba$ . È così dimostrato che  $R$  è commutativo.  $\square$

Enunciamo ma non dimostriamo il teorema di Stone, che è il risultato fondamentale nella teoria degli anelli booleani.

**Teorema di Stone.** *Sia  $R$  un anello booleano. Allora:*

- (i) esiste un insieme  $S$  tale che  $R$  sia isomorfo ad un sottoanello unitario di  $(\mathcal{P}(S), \Delta, \cap)$ ;
- (ii) se  $R$  è finito, esiste un insieme  $S$  tale che  $R$  sia isomorfo a  $(\mathcal{P}(S), \Delta, \cap)$ .

Va notato, a proposito del punto (i), che tutti i sottoanelli unitari di  $(\mathcal{P}(S), \Delta, \cap)$  sono booleani. Infatti:

**Esercizio 2.** Se  $R$  è un anello booleano ogni sottoanello unitario di  $R$  è booleano.

Il teorema di Stone ha un'importante conseguenza:

**Corollario 3.** *Sia  $R$  un anello booleano finito. Allora:*

- (i)  $|R|$  è un potenza di 2;
- (ii) se  $A$  è un anello booleano e  $|A| = |R|$ , allora  $A \simeq R$ .

*Dimostrazione.* Per il teorema di Stone, esiste un insieme  $S$ , ovviamente finito, tale che  $R$  sia isomorfo a  $(\mathcal{P}(S), \Delta, \cap)$ . Posto  $n = |S|$ , allora  $|R| = |\mathcal{P}(S)| = 2^n$ , il che giustifica la (i). Se poi  $A$  è un anello booleano, anch'esso di cardinalità  $2^n$ , ancora per il teorema di Stone abbiamo  $A \simeq (\mathcal{P}(T), \Delta, \cap)$  per un opportuno insieme  $T$ . Ma allora  $|\mathcal{P}(T)| = |A|$ , quindi  $|\mathcal{P}(T)| = 2^n$  e deduciamo così  $|T| = n$ . Dunque,  $|T| = |S|$ ; questo comporta (vedi l'esercizio che segue)  $(\mathcal{P}(T), \Delta, \cap) \simeq (\mathcal{P}(S), \Delta, \cap)$ , quindi  $A \simeq R$ .  $\square$

**Esercizio 4.** Verificare che se  $f: S \rightarrow T$  è un'applicazione biettiva, allora l'applicazione immagine  $\vec{f}: \mathcal{P}(S) \rightarrow \mathcal{P}(T)$  è un isomorfismo di anelli da  $(\mathcal{P}(S), \Delta, \cap)$  a  $(\mathcal{P}(T), \Delta, \cap)$ .

È bene notare che, nel teorema di Stone il caso degli anelli booleani infiniti differisce effettivamente dal caso degli anelli finiti. Esistono infatti anelli booleani infiniti che non sono isomorfi a  $(\mathcal{P}(S), \Delta, \cap)$  per alcun insieme  $S$ . Un esempio è fornito dall'insieme  $P$  costituito da tutti i sottoinsiemi  $X$  di  $\mathbb{N}$  tali che uno tra  $X$  e  $\mathbb{N} \setminus X$  sia finito.<sup>(2)</sup> Non è difficile verificare (ed è un buon esercizio farlo) che  $P$  è un sottoanello unitario di  $(\mathcal{P}(\mathbb{N}), \Delta, \cap)$  e di conseguenza è un anello booleano. Si può però dimostrare (ma non si tratta in questo caso di un esercizio) che per ogni insieme  $S$  non esistono applicazioni biettive da  $P$  a  $\mathcal{P}(S)$ , quindi  $P$ , come anello, non può essere isomorfo a  $(\mathcal{P}(S), \Delta, \cap)$ .

<sup>(1)</sup>Se invece non esiste nessun  $n \in \mathbb{N}^*$ , tale che  $n1_R = 0_R$ , cioè: se  $1_R$  non è periodico in  $(R, +)$ , allora  $R$  ha per definizione caratteristica 0.

<sup>(2)</sup>una parte  $X$  di un insieme  $Y$  si dice cofinita in  $Y$  se e solo se  $Y \setminus X$  è un insieme finito. Dunque,  $P$  è l'insieme costituito dalle parti finite e dalle parti cofinite di  $\mathbb{N}$ .

## 2. RETICOLI BOOLEANI

Ricordiamo<sup>(3)</sup> che un reticolo è un insieme ordinato non vuoto  $(L, \leq)$  tale che, per ogni  $a, b \in L$  esistano l'estremo inferiore  $\inf_{(L, \leq)}\{a, b\}$  e l'estremo superiore  $\sup_{(L, \leq)}\{a, b\}$  di  $\{a, b\}$  in  $(L, \leq)$ .

Ricordiamo anche che si può, in modo equivalente, riguardare i reticoli anche come strutture algebriche. Infatti, se  $(L, \leq)$  è un reticolo, si definiscono in  $L$  le due *operazioni reticolari*  $\vee$  e  $\wedge$ , ponendo, per ogni  $a, b \in L$ ,

$$a \vee b = \sup_{(L, \leq)}\{a, b\} \quad \text{e} \quad a \wedge b = \inf_{(L, \leq)}\{a, b\}$$

e valgono, per  $\vee$  e  $\wedge$  queste proprietà algebriche:

- (1)  $\vee$  e  $\wedge$  sono commutative;
- (2)  $\vee$  e  $\wedge$  sono associative;
- (3) valgono le leggi di assorbimento: per ogni  $a, b \in L$ ,
  - $a \vee (a \wedge b) = a$ ;
  - $a \wedge (a \vee b) = a$ .

Vale anche per  $\vee$  e  $\wedge$  una quarta proprietà, l'iteratività: per ogni  $a \in L$ ,  $a \vee a = a = a \wedge a$  (vale a dire: ogni elemento di  $L$  è idempotente sia rispetto a  $\vee$  che rispetto a  $\wedge$ ). Se, viceversa,  $(L, \vee, \wedge)$  è una struttura algebrica in cui  $\vee$  e  $\wedge$  sono due operazioni binarie che verificano (1), (2) e (3), allora si può definire in  $L$  una relazione binaria  $\preceq$  ponendo, per ogni  $a, b \in L$ ,

$$a \preceq b \iff a = a \wedge b$$

e si verifica che  $\preceq$  è una relazione d'ordine che rende  $(L, \preceq)$  un reticolo. Inoltre, per ogni  $a, b \in L$  si ha  $a \vee b = \sup_{(L, \preceq)}\{a, b\}$  e  $a \wedge b = \inf_{(L, \preceq)}\{a, b\}$ . Dunque,  $\vee$  e  $\wedge$  risultano essere le operazioni reticolari in  $(L, \preceq)$ . Allo stesso modo, se  $\vee$  e  $\wedge$  sono le operazioni reticolari definite in un reticolo  $(L, \leq)$ , è chiaro che la relazione  $\preceq$  definita sopra coincide con  $\leq$ .

In sintesi, fissato un insieme non vuoto  $L$ , se  $\mathcal{A}$  è l'insieme delle relazioni d'ordine  $\leq$  tali che  $(L, \leq)$  sia un reticolo e  $\mathcal{B}$  è l'insieme delle coppie  $(\vee, \wedge)$  di operazioni binarie in  $L$  che verificano le condizioni (1), (2) e (3), abbiamo definito due applicazioni tra  $\mathcal{A}$  e  $\mathcal{B}$ . La prima è  $\alpha: \mathcal{A} \rightarrow \mathcal{B}$ , che ad una relazione d'ordine  $\leq \in \mathcal{A}$  associa la coppia ordinata  $(\vee, \wedge) \in \mathcal{B}$ , dove  $\vee$  e  $\wedge$  sono le operazioni reticolari di estremo superiore ed estremo inferiore in  $(L, \leq)$ . La seconda applicazione è  $\beta: \mathcal{B} \rightarrow \mathcal{A}$ , che ad ogni  $(\vee, \wedge) \in \mathcal{B}$  associa la relazione d'ordine  $\preceq \in \mathcal{A}$  definita come sopra. Quello che abbiamo evidenziato è che  $\alpha$  e  $\beta$  sono l'una inversa dell'altra, quindi sono biettive.

L'esistenza di queste biezioni fa sì che sia del tutto equivalente lo studio dei reticoli (intesi come particolari insiemi ordinati) e quello delle strutture algebriche  $(L, \vee, \wedge)$  per le quali valgano le condizioni (1), (2) e (3). Per questo motivo si fa riferimento a queste strutture chiamandole 'reticoli come strutture algebriche'. D'ora in avanti, dunque, per indicare un reticolo faremo indifferentemente riferimento alla struttura di insieme ordinato (indicando, ad esempio, il reticolo come  $(L, \leq)$ ) o alla struttura algebrica (indicando il reticolo, con un abuso di terminologia, come  $(L, \vee, \wedge)$ ); conveniamo che la prima operazione indicata è quella di estremo superiore, la seconda quella di estremo inferiore). Può essere conveniente, e lo faremo, indicare un reticolo come  $(L, \leq, \vee, \wedge)$  per specificare in modo sintetico sia la relazione d'ordine che le operazioni reticolari.

Ricordiamo che anche le due possibili nozioni di isomorfismo per i reticoli (come insiemi ordinati) ed i reticoli come strutture algebriche coincidono. Va però osservato che la nozione di *sottoreticolo* è algebrica, nel senso che può essere definita solo in termini delle operazioni reticolari.

Infatti, se  $(L, \leq)$  è un reticolo, un sottoreticolo di  $(L, \leq)$  è per definizione un sottoinsieme non vuoto  $K$  di  $L$  che sia chiuso rispetto alle operazioni reticolari  $\vee$  e  $\wedge$  di  $L$ . Le operazioni indotte in  $K$  da  $\vee$  e  $\wedge$  continuano a verificare le condizioni (1), (2) e (3) e quindi rendono  $K$  un reticolo rispetto alla relazione d'ordine indotta da  $\leq$  su  $K$  (quest'ultima osservazione è garantita dal fatto che la relazione d'ordine del reticolo è determinata dalle operazioni reticolari: per ogni  $a, b \in L$  si ha  $a \leq b \iff a = a \wedge b$ ).

Se  $a$  e  $b$  sono elementi di un insieme ordinato  $(S, \leq)$ , si chiama intervallo chiuso di estremi  $a$  e  $b$ , e si indica con  $[a, b]_{(S, \leq)}$  (o semplicemente  $[a, b]$  se il riferimento a  $(S, \leq)$  può essere sottinteso) l'insieme  $\{x \in S \mid a \leq x \leq b\}$ , che è diverso dal vuoto se e solo se  $a \leq b$  e in questo caso ha  $a$  come minimo e  $b$  come massimo.

**Lemma 5.** Siano  $a$  e  $b$  elementi del reticolo  $(L, \leq)$ .

- (i) l'insieme  $\{x \in L \mid a \leq x\}$  è un sottoreticolo di  $(L, \leq)$ ;
- (ii) l'insieme  $\{x \in L \mid x \leq b\}$  è un sottoreticolo di  $(L, \leq)$ ;
- (iii) se  $a \leq b$ , l'intervallo chiuso  $[a, b]$  è un sottoreticolo di  $(L, \leq)$ .

*Dimostrazione.* Sia  $X = \{x \in L \mid a \leq x\}$ . Certamente  $X \neq \emptyset$ , perché  $a \in X$ . Siano  $x$  e  $y$  elementi di  $X$ . Allora  $a \leq x$  e  $a \leq y$ , quindi  $a$  è un minorante di  $\{x, y\}$  in  $(L, \leq)$ . Dunque  $a \leq \inf_{(L, \leq)}\{x, y\} = x \wedge y$  e possiamo concludere  $x \wedge y \in X$ . Inoltre  $a \leq x \leq x \vee y$ , quindi  $x \vee y \in X$ . Abbiamo così provato che  $X$  è chiuso rispetto a  $\vee$  e  $\wedge$ , quindi è un sottoreticolo di  $(L, \leq)$ . È così provata la (i). Per dualità, anche  $Y := \{x \in L \mid x \leq b\}$  è un sottoreticolo, quindi anche la (ii) è vera. Infine, si ha ovviamente  $[a, b] = X \cap Y$ , quindi  $[a, b]$  è chiuso rispetto a  $\vee$  e  $\wedge$ , in quanto intersezione di parti chiuse. Se  $a \leq b$ , allora  $[a, b] \neq \emptyset$  e  $[a, b]$  è un sottoreticolo di  $L$ ; vale così anche (iii).  $\square$

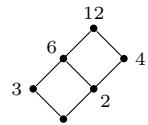
---

<sup>(3)</sup>per tutto ciò che qui viene 'ricordato' e non giustificato o comunque spiegato in dettaglio, si rimanda al libro di testo o alle altre fonti a disposizione.

Ad esempio, per ogni  $n \in \mathbb{N}$  sia l'insieme  $\text{Div}_{\mathbb{N}}(n)$  dei divisori di  $n$  che quello,  $n\mathbb{N}$ , dei multipli di  $n$  (in  $\mathbb{N}$ ) costituiscono sottoreticolari di  $(\mathbb{N}, |)$ . Similmente, se  $S$  è un insieme e  $T \subseteq S$ , allora sia  $\mathcal{P}(T)$  che l'insieme delle parti di  $S$  contenenti  $T$  costituiscono sottoreticolari di  $(\mathcal{P}(S), \subseteq)$ .

**Esempio 6.** Sia  $L = \text{Div}_{\mathbb{N}}(12)$  il reticolo dei divisori di 12, rappresentato dal diagramma di Hasse a destra. Il sottoinsieme  $K = L \setminus \{6\}$  non è un sottoreticolo di  $L$ , infatti  $K$  non è chiuso

rispetto all'operazione reticolare  $\vee$ , dal momento che 2 e 3 appartengono a  $K$  ma  $6 = 2 \vee 3 \notin K$ . Se però consideriamo  $K$  come insieme ordinato dall'ordinamento indotto da quello di  $L$ , quindi ordinato per divisibilità, non è difficile verificare che, rispetto a questo ordinamento,  $K$  è un reticolo; il suo diagramma di Hasse è rappresentato a sinistra.



Questo esempio mostra che anche se un sottoinsieme ordinato di un reticolo  $L$  è, rispetto all'ordinamento indotto, a sua volta un reticolo, non è detto che esso sia un sottoreticolo di  $L$ .

Anche le nozioni di minimo e massimo hanno un'interpretazione algebrica.

**Lemma 7.** Sia  $(L, \leq, \vee, \wedge)$ , un reticolo. Per ogni  $a \in L$ ,  $a$  è il minimo in  $L$  se e solo se  $a$  è elemento neutro rispetto a  $\vee$ ;  $a$  è il massimo in  $L$  se e solo se  $a$  è elemento neutro rispetto a  $\wedge$ .

*Dimostrazione.* Si ha  $a = \min L$  se e solo se  $a \leq b$  per ogni  $b \in L$ ; ma  $a \leq b$  equivale a  $a \vee b = b$ . Dunque,  $a = \min L$  se e solo se, per ogni  $b \in L$  si ha  $a \vee b = b$ , cioè: se e solo se  $a$  è neutro in  $(L, \vee)$ . È così provata la prima parte dell'enunciato. La seconda è duale.  $\square$

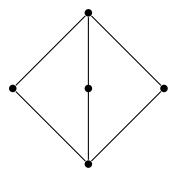
Dunque, se  $(L, \leq, \vee, \wedge)$  è un reticolo limitato (cioè dotato di minimo e massimo) sia  $(L, \vee)$  che  $(L, \wedge)$  sono monoidi commutativi.

Ad esempio, il reticolo  $(\mathcal{P}(S), \subseteq)$  delle parti di un insieme  $S$  ha minimo e massimo, rispettivamente  $\emptyset$  e  $S$ , ed operazioni reticolari  $\cup$  e  $\cap$ . In effetti,  $\emptyset$  è l'elemento neutro del monoide  $(S, \cup)$ ,  $S$  è l'elemento neutro del monoide  $(S, \cap)$ .

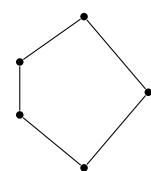
**Definizione.** Sia  $(L, \leq, \vee, \wedge)$  un reticolo limitato e sia  $a \in L$ . Per ogni  $b \in L$ ,  $b$  è un *complemento* di  $a$  in  $L$  se e solo se  $a \vee b = \max L$  e  $a \wedge b = \min L$ .

Dovrebbe essere chiaro che, con le notazioni della definizione, dire che  $b$  è un complemento di  $a$  equivale a dire che  $a$  è un complemento di  $b$ . Altrettanto ovvio è che  $\min L$  e  $\max L$  sono l'uno complemento dell'altro (anzi,  $\min L$  è l'unico complemento di  $\max L$  in  $L$  e, dualmente,  $\max L$  è l'unico complemento di  $\min L$  in  $L$ .) Altri esempi:

- nel reticolo  $(\mathcal{P}(S), \subseteq)$  delle parti di un insieme  $S$ , ogni elemento ha uno ed un solo complemento. Infatti, per ogni  $X \in \mathcal{P}(S)$ ,  $X \cup (S \setminus X) = S = \max \mathcal{P}(S)$  e  $X \cap (S \setminus X) = \emptyset = \min \mathcal{P}(S)$ , quindi  $S \setminus X$  è un complemento di  $X$  in  $\mathcal{P}(S)$ . L'unicità è facile da verificare direttamente, ma segue anche da considerazioni che faremo più avanti ([Proposizione 9](#)).
- Nel reticolo dei divisori di 12, visto nell'[Esempio 6](#), gli elementi 1 e 12 (minimo e massimo del reticolo) sono l'uno complemento dell'altro, 3 e 4 sono l'uno complemento dell'altro ma né 2 né 6 hanno complemento.
- Come si vede facilmente, se  $L$  è un insieme non vuoto totalmente ordinato (e quindi un reticolo) limitato, in  $L$  gli unici elementi che hanno complemento sono il minimo ed il massimo.
- Anche in  $(\mathbb{N}, |)$ , gli unici elementi che hanno complemento sono il minimo, 0, ed il massimo, 1. Sia infatti  $a \in \mathbb{N}$  e sia  $b$  un complemento di  $a$  in  $(\mathbb{N}, |)$ . Ricordando che le operazioni reticolari in  $(\mathbb{N}, |)$  sono descritte dal minimo comune multiplo e dal massimo comun divisore, abbiamo  $0 = \text{mcm}(a, b)$  e  $1 = \text{MCD}(a, b)$ . Se  $a \neq 0$ , allora da  $\text{mcm}(a, b) = 0$  segue  $b = 0$ , ma se  $b = 0$  allora  $1 = \text{MCD}(a, b) = \text{MCD}(a, 0) = a$ . Dunque, se  $a \notin \{0, 1\}$ ,  $a$  non ha complementi in  $(\mathbb{N}, |)$ .
- Un elemento in un reticolo (limitato) può anche avere più di un complemento. Questi due esempi sono di grande importanza:



reticolo trirettangolo



reticolo pentagonale

Come si vede immediatamente, nel reticolo trirettangolo ciascuno dei tre elementi diversi dal minimo e dal massimo ha gli altri due come complementi; nel reticolo pentagonale l'elemento rappresentato più a destra ha due complementi.

Ovviamente è possibile modificare questi esempi in modo da ottenere reticolli finiti (e quindi limitati) con elementi dotati di un numero arbitrario di complementi. (Come?)

**Definizione.** Un reticolo  $L$  si dice *complementato* se e solo se ogni suo elemento ha in  $L$  almeno un complemento.

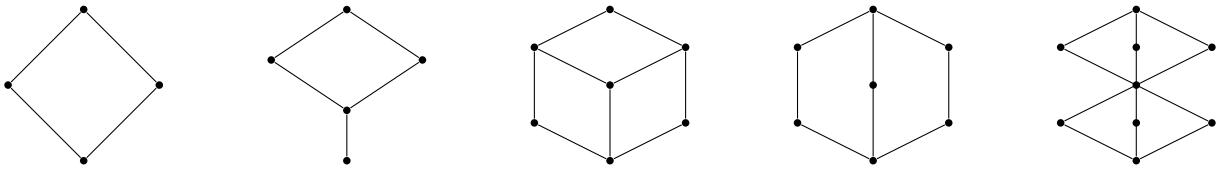
È chiaro che un reticolo, per essere complementato deve essere, in primo luogo, limitato, altrimenti in esso non possono esistere elementi dotati di complementi. Dagli esempi forniti a proposito della nozione di complemento vediamo subito che:

- per ogni insieme  $S$ , il reticolo  $(\mathcal{P}(S), \subseteq)$  è complementato;
- un insieme non vuoto totalmente ordinato è complementato se e solo se ha al massimo due elementi;
- né  $(\mathbb{N}, |)$  né il reticolo dei divisori di 12 sono complementati;
- il reticolo trirettangolo e quello pentagonale sono complementati.

**Esercizio 8.** Per ogni  $n \in \mathbb{N}$ , sia  $D_n$  il reticolo dei divisori di  $n$  in  $\mathbb{N}$  (che è, ricordiamo, un sottoreticolo di  $(\mathbb{N}, |)$ ). Lo scopo di questo esercizio è riconoscere che  $D_n$  è complementato se e solo se  $n$  è un intero *libero da quadrati*, cioè un intero non divisibile per il quadrato di alcun primo.<sup>(4)</sup>

- Sia  $d$  un divisore (in  $\mathbb{N}$ ) di  $n$ . Se  $d$  e  $n/d$  sono coprimi, allora  $n/d$  è un complemento di  $d$  in  $D_n$ . [Suggerimento: basta calcolare MCD e mcm tra  $d$  e  $n/d$ .]
- Dedurre dal punto precedente che se  $n$  è libero da quadrati allora  $D_n$  è complementato. [Suggerimento: pensare alla scomposizione di  $n$  in fattori primi e descrivere i divisori di  $n$ .]
- Supponiamo che esista un primo  $p$  tale che  $p^2$  divida  $n$ . Allora  $p$  non ha complemento in  $D_n$ . [Suggerimento: se  $a$  è un complemento di  $p$ ,  $p$  divide o non divide  $a$ ?]
- A questo punto la conclusione è facile:  $D_n$  è complementato se e solo se  $n$  è libero da quadrati.

Ulteriori esempi: dei reticolari qui rappresentati sono complementati il primo, ed il quarto, non gli altri tre.



Un'altra proprietà di natura algebrica riferita a reticolari è la distributività.

**Definizione.** Un reticolo  $(L, \leq, \vee, \wedge)$  si dice *distributivo* se e solo ciascuna delle due operazioni reticolari  $\vee$  e  $\wedge$  è distributiva rispetto all'altra.

In termini più esplicativi,  $(L, \leq, \vee, \wedge)$  è distributivo se e solo se, per ogni  $a, b, c \in L$  si ha:

- $$(d_1): a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c); \text{ e}$$
- $$(d_2): a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c).$$

In realtà è possibile dimostrare che se, in un reticolo  $L$ , è verificata almeno una delle due condizioni  $(d_1)$  e  $(d_2)$  per ogni terna  $(a, b, c)$  di elementi di  $L$ , allora anche l'altra è verificata e quindi  $L$  è distributivo.

Ad esempio, l'operazione insiemistica di unione binaria è distributiva rispetto all'intersezione, e viceversa l'intersezione è distributiva rispetto all'unione, quindi, per ogni insieme  $S$ , il reticolo  $(\mathcal{P}(S), \subseteq)$  è distributivo.

Non è difficile verificare (è un utile esercizio di aritmetica) che anche il reticolo  $(\mathbb{N}, |)$  è distributivo, così come sono distributivi i reticolari totalmente ordinati (quest'ultimo fatto segue anche dal criterio di distributività di Birkhoff, che incontreremo tra poco).

Invece, non sono distributivi né il reticolo trirettangolo né il reticolo pentagonale. Questo fatto segue dal prossimo risultato, perché come abbiamo visto, in questi due reticolari esistono elementi con più complementi.

**Proposizione 9.** Sia  $(L, \leq, \vee, \wedge)$  un reticolo distributivo. Allora ogni elemento di  $L$  ha al più un complemento in  $L$ .

*Dimostrazione.* Sia  $a \in L$ , e siano  $x$  e  $y$  complementi di  $a$  in  $L$ . Per provare l'enunciato occorre (e basta) verificare che  $x = y$ .

Indicando con 1 e 0, nell'ordine, il massimo e il minimo di  $L$ , si ha  $a \wedge x = a \wedge y = 0$  e  $a \vee x = a \vee y = 1$ . Usando la proprietà distributiva abbiamo:

$$x = x \vee 0 = x \vee (a \wedge y) = (x \vee a) \wedge (x \vee y) = 1 \wedge (x \vee y) = x \vee y.$$

Analogamente, scambiando i ruoli tra  $x$  e  $y$ , possiamo ottenere  $y = y \vee x = x \vee y = x$ .  $\square$

Dovrebbe essere evidente dalla definizione che ogni sottoreticolo di un reticolo distributivo è a sua volta distributivo. Di conseguenza, un reticolo distributivo non può avere sottoreticolari che siano isomorfi al reticolo trirettangolo o quello pentagonale. Un risultato notevole della teoria dei reticolari, che non dimostreremo, mostra che vale anche il viceversa: l'assenza di tali sottoreticolari basta a provare che un reticolo è distributivo.

**Criterio di distributività di Birkhoff.** Sia  $L$  un reticolo.  $L$  è distributivo se e solo se non ha sottoreticolari isomorfi a uno tra il reticolo trirettangolo e il reticolo pentagonale.

<sup>(4)</sup>i numeri naturali liberi da quadrati sono dunque 1 ed i numeri naturali che si possono scrivere come prodotti di primi a due a due distinti.

Vediamo qualche esempio di applicazione del criterio di Birkhoff. Poiché i sottoreticolati dei reticolati totalmente ordinati sono certamente totalmente ordinati, e nessuno dei due reticolati trirettangolo o pentagonale lo è, il criterio di Birkhoff fornisce una maniera per dimostrare che i reticolati totalmente ordinati sono distributivi. Siccome sia il reticolo trirettangolo che quello pentagonale sono costituiti da cinque elementi, il criterio di Birkhoff mostra anche che i reticolati con meno di cinque elementi sono sicuramente distributivi, quelli di cardinalità cinque sono distributivi se e solo se non sono isomorfi né al reticolo trirettangolo né al pentagonale. Se torniamo ai cinque reticolati esaminati come esempi dopo l'[Esercizio 8](#), vediamo così che i primi due sono distributivi, gli altri tre no. Evidenziamo sottoreticolati trirettangoli (in blu) o pentagonali (in rosso) nei tre reticolati non distributivi:



È necessario fare attenzione al fatto che il criterio di Birkhoff esclude l'esistenza di sottoreticolati isomorfi al reticolo trirettangolo o a quello pentagonale in un reticolo distributivo  $L$ , ma non esclude che un sottoinsieme di  $L$ , munito dell'ordinamento indotto, possa essere un reticolo di uno di questi due tipi. Consideriamo il reticolo  $L$  dei divisori di 12 ed il suo sottoinsieme  $K = L \setminus \{6\}$  discusso nell'[Esempio 6](#). Come sappiamo,  $L$  è un reticolo distributivo (è un sottoreticolo di  $(\mathbb{N}, |)$ ) e, come si può vedere,  $K$  è isomorfo al reticolo pentagonale. Questo non contraddice il criterio di Birkhoff, perché  $K$  non è un sottoreticolo di  $L$ .

**Definizione.** Un reticolo si dice *booleano* se e solo se è distributivo e complementato.

Ad esempio, per ogni insieme  $S$ , il reticolo  $(\mathcal{P}(S), \subseteq)$  è booleano. In conseguenza della definizione e della [Proposizione 9](#), se  $L$  è un reticolo booleano, ogni elemento di  $L$  ha uno ed un solo complemento in  $L$ .

Osserviamo che un reticolo  $L$  è complementato, distributivo o booleano, allora anche il duale di  $L$  ha la stessa proprietà. Quindi vale per i reticolati con queste proprietà il principio di dualità: se una certa affermazione è verificata da ogni reticolo complementato, allora anche l'affermazione duale varrà in ogni reticolo complementato. Lo stesso è vero se nella frase precedente sostituiamo “complementato” con “distributivo” o con “booleano”.

### 3. ALGEBRE DI BOOLE

Come sappiamo, si può dare la nozione di reticolo in termini puramente algebrici, cioè esclusivamente in termini di operazioni, senza fare riferimento a relazioni d'ordine: stiamo parlando dei reticolati ‘come strutture algebriche’. Sia  $(L, \wedge, \vee)$  una struttura algebrica di reticolo; vediamo quali condizioni sulle operazioni dobbiamo imporre affinché il reticolo  $L$  sia booleano. Oltre alle proprietà commutativa, associativa ed alle leggi di assorbimento, che già conosciamo, devono valere le proprietà distributive (di  $\vee$  rispetto a  $\wedge$  e di  $\wedge$  rispetto a  $\vee$ ), che fanno sì che il reticolo  $L$  sia distributivo. Sappiamo poi dal [Lemma 7](#) che il fatto che  $L$  sia limitato equivale all'esistenza di elementi neutri per  $\vee$  e  $\wedge$ . Infine, come abbiamo visto, in un reticolo booleano ogni elemento ha un unico complemento; possiamo allora considerare l'applicazione  $'$ :  $L \rightarrow L$  che ad ogni  $a \in L$  associa il suo complemento  $a'$  in  $L$ . Queste considerazioni suggeriscono la seguente definizione:

**Definizione.** Si dice *algebra di Boole* una struttura algebrica  $(L, \vee, \wedge, 0, 1, ')$ , dove  $\vee$  e  $\wedge$  sono operazioni binarie, 0 e 1 operazioni nullarie e  $'$  un'operazione unaria, tale che:

- (1)  $(L, \vee, 0)$  e  $(L, \wedge, 1)$  siano monoidi commutativi;
- (2) valgano le leggi di assorbimento: per ogni  $a, b \in L$ ,  $a \vee (a \wedge b) = a = a \wedge (a \vee b)$ ;
- (3)  $\vee$  sia distributiva rispetto a  $\wedge$  e  $\wedge$  sia distributiva rispetto a  $\vee$ ;
- (4) per ogni  $a \in L$ ,  $a \vee a' = 1$  e  $a \wedge a' = 0$ ,

dove abbiamo indicato con  $a'$  l'immagine di  $a$  rispetto a  $'$ .

Per quanto detto sopra, ogni reticolo booleano dà luogo ad un'algebra di Boole, viceversa un'algebra di Boole si può sempre riguardare come reticolo booleano. Infatti, la (1) e la (2) esprimono esattamente il fatto che  $(L, \wedge, \vee)$  è un reticolo limitato, con minimo 0 e massimo 1, come segue dal [Lemma 7](#); la (3) dice che questo reticolo è distributivo e la (4) garantisce che ogni elemento  $a$  di  $L$  ha un complemento:  $a'$ .

Possiamo dunque dire che la nozione di algebra di Boole è la versione ‘puramente algebrica’ della nozione di reticolo booleano.

Abbiamo, come per tutti tipi di strutture algebriche, una nozione di isomorfismo tra algebre di Boole: un'isomorfismo da un'algebra di Boole  $(L_1, \vee_1, \wedge_1, 0_1, 1_1, ')$  ad un'algebra di Boole  $(L_2, \vee_2, \wedge_2, 0_2, 1_2, '')$  è un'applicazione biettiva  $f: L_1 \rightarrow L_2$  che ‘conservi le operazioni’, tale cioè che, per ogni  $a, b \in L_1$  si abbia

- i.)  $f(a \vee_1 b) = f(a) \vee_2 f(b)$  e  $f(a \wedge_1 b) = f(a) \wedge_2 f(b)$ ;
- ii.)  $f(0_1) = 0_2$  e  $f(1_1) = 1_2$ ;
- iii.)  $f(a') = (f(a))''$ .

Ora, la i.), cioè il proprietà che  $f$  conservi le operazioni reticolari, equivale al fatto che la biezione  $f$  sia un isomorfismo di reticolati. Se questa proprietà è verificata valgono però anche la ii.) e la iii.). Infatti, se  $f$  è un isomorfismo di reticolati da  $L_1$  a  $L_2$ , allora  $f$  deve mandare il minimo  $0_1$  di  $L_1$  nel minimo  $0_2$  di  $L_2$  e, analogamente,  $1_1 = \max L_1$  in  $1_2 = \max L_2$ . Vale così la ii.). Inoltre, per ogni  $a \in L_1$ , poiché  $a'$  è un complemento di  $a$  in  $L_1$ ,

la sua immagine  $f(a')$  deve essere un complemento di  $f(a)$  in  $L_2$ . Ma, poiché  $L_2$  è booleano,  $(f(a))''$  è l'unico complemento di  $f(a)$  in  $L_2$ , quindi  $f(a') = (f(a))''$ . Quello che abbiamo verificato è che gli isomorfismi di algebre di Boole da  $L_1$  a  $L_2$  sono tutti e soli gli isomorfismi di reticolati da  $L_1$  a  $L_2$ . In particolare due algebre di Boole sono isomorfe (come algebre di Boole) se e solo se sono isomorfe come reticolati. A questo punto possiamo davvero concludere che lo studio delle algebre di Boole equivale allo studio dei reticolati booleani.

**Definizione.** Sia  $(L, \vee, \wedge, 0, 1, ')$  un'algebra di Boole. Una parte non vuota  $K$  di  $L$  ne costituisce una *sottoalgebra di Boole* se e solo se  $K$  è un sottomonoide sia di  $(L, \vee)$  che di  $(L, \wedge)$  e contiene il complemento in  $L$  di ogni suo elemento.

In modo più esplicito,  $K$  costituisce una sottoalgebra di Boole di  $(L, \vee, \wedge, 0, 1, ')$  se e solo se  $K \subseteq L$  e sono verificate queste condizioni: per ogni  $a, b \in K$ ,

- $a \vee b \in K$  e  $a \wedge b \in K$ ;
- $0 \in K$  e  $1 \in K$ ;
- $a' \in K$ .

È evidente che in queste condizioni  $K$ , munita delle operazioni indotte da quelle di  $L$  è a sua volta un'algebra di Boole.

La nozione di sottoalgebra di Boole differisce da quella di sottoreticolo. Infatti, un sottoreticolo  $K$  di un reticolo booleano  $L$  deve essere chiuso rispetto alle due operazioni reticolari (quindi deve verificare la prima delle tre condizioni appena elencate), ma non contiene necessariamente il massimo o il minimo del reticolo né, tanto meno, i complementi dei suoi elementi.

**Esempio 10.** Dato un insieme  $S \neq \emptyset$ , consideriamo il reticolo booleano  $(\mathcal{P}(S), \subseteq)$ . Questo si struttura come algebra di Boole nella forma  $(\mathcal{P}(S), \cup, \cap, \emptyset, S, ^c)$ , dove  $^c$  è l'applicazione “complemento” che manda ogni  $X \in \mathcal{P}(S)$  in  $X^c = S \setminus X \in \mathcal{P}(S)$ . Se  $T$  è una parte propria di  $S$ , allora  $\mathcal{P}(T)$  costituisce un sottoreticolo di  $\mathcal{P}(S)$  (ad esempio, per il [Lemma 5](#)), ma non una sottoalgebra di Boole di  $\mathcal{P}(S)$ , dal momento che  $S \notin \mathcal{P}(T)$ .

Nel caso appena considerato,  $(\mathcal{P}(T), \subseteq)$  è comunque un reticolo booleano, quindi si struttura come algebra di Boole. Ma in altri casi la situazione può essere diversa. Ad esempio, se supponiamo  $\emptyset \neq T \subset S$ , allora  $\{\emptyset, T, S\}$  forma un sottoreticolo di  $(\mathcal{P}(S), \subseteq)$  che non è complementato e quindi non è booleano.

**Esercizio 11.** Provare che un parte di un'algebra di Boole ne è una sottoalgebra di Boole se e solo se è un sottoreticolo che contenga il complemento di ogni suo elemento.

Il prossimo enunciato elenca alcune identità che valgono nelle algebre di Boole. La terza si esprime dicendo che l'operazione di complemento è involutoria, cioè coincide con l'applicazione inversa di sé stessa (e, in particolare, è biettiva); le ultime due sono le note come leggi di De Morgan per algebre di Boole.

**Proposizione 12.** Sia  $(L, \vee, \wedge, 0, 1, ')$  un'algebra di Boole. Allora, per ogni  $a, b \in L$ ,

- (i)  $1 \vee a = 1$  e  $0 \wedge a = 0$ ;
- (ii)  $1' = 0$  e  $0' = 1$ ;
- (iii)  $(a')' = a$ ;
- (iv)  $(a \vee b)' = a' \wedge b'$ ;
- (v)  $(a \wedge b)' = a' \vee b'$ .

*Dimostrazione.* La (i) e la (ii) sono immediate: visto  $L$  come reticolo, e quindi come insieme ordinato, 1 e 0 ne sono il massimo e il minimo e le operazioni  $\vee$  e  $\wedge$  forniscono estremi superiori e inferiori, dunque  $1 \vee a = \sup\{1, a\} = 1$  e  $0 \wedge a = \inf\{0, a\} = 0$ ; (5) inoltre, come sappiamo, minimo e massimo sono sempre l'uno il complemento dell'altro.

Anche la (iii) è pressoché ovvia: essendo  $a'$  un complemento di  $a$ ,  $a$  è un complemento di  $a'$ . Anche  $(a')'$  è un complemento di  $a'$ ; l'unicità dei complementi nei reticolati booleani comporta  $a = (a')'$ .

Sempre per l'unicità del complemento, per provare la (iv) basterà mostrare che  $a' \wedge b'$  è un complemento di  $a \vee b$ , cioè  $(a \vee b) \vee (a' \wedge b') = 1$  e  $(a \vee b) \wedge (a' \wedge b') = 0$ . Usando la distributività e la (i), abbiamo  $(a \vee b) \vee (a' \wedge b') = (a \vee b \vee a') \wedge (a \vee b \vee b') = (a \vee a' \vee b) \wedge (a \vee b \vee b') = (1 \vee b) \wedge (a \vee 1) = 1 \wedge 1 = 1$  e, in modo simmetrico,  $(a \vee b) \wedge (a' \wedge b') = (a \wedge a' \wedge b') \vee (b \wedge a' \wedge b') = (a \wedge a' \wedge b') \vee (b \wedge b' \wedge a') = (0 \wedge b') \vee (0 \wedge a') = 0 \vee 0 = 0$ . È così provata la (iv). La (v) segue per dualità.  $\square$

Ad illustrazione della [Proposizione 12](#), leggiamo le identità appena provate nel caso in cui l'algebra di Boole  $L$  che appare nell'enunciato sia l'algebra delle parti di un insieme  $S$ , cioè  $(\mathcal{P}(S), \cup, \cap, \emptyset, S, ^c)$ , dove, come nell'[Esempio 10](#):  $X \in \mathcal{P}(S) \mapsto S \setminus X \in \mathcal{P}(S)$ . In questo caso la [Proposizione 12](#) esprime cinque ben note formule insiemistiche elementari: per ogni  $a, b \in \mathcal{P}(S)$ , (i):  $S \cup a = S$  e  $\emptyset \cap a = \emptyset$ ; (ii):  $S \setminus S = \emptyset$  e  $S \setminus \emptyset = S$ ; (iii):  $S \setminus (S \setminus a) = a$ ; (iv):  $S \setminus (a \cup b) = (S \setminus a) \cap (S \setminus b)$  e (v):  $S \setminus (a \cap b) = (S \setminus a) \cup (S \setminus b)$ .

**Esercizio 13.** Sia  $(L, \vee, \wedge, 0, 1, ')$  un'algebra di Boole. Allora anche  $(L, \wedge, \vee, 1, 0, ')$  è un'algebra di Boole, quella costruita a partire dal reticolo booleano  $(L, \wedge, \vee)$ , il duale di  $(L, \vee, \wedge)$ . Verificare che l'applicazione  $'$  è un isomorfismo tra queste due algebre di Boole. Questa è una riformulazione della [Proposizione 12](#). Notare la conseguenza: ogni reticolo booleano è isomorfo al suo duale.

(5)oppure, per via algebrica: dal momento che 1 è neutro rispetto a  $\wedge$ ,  $1 \wedge a = a$ , quindi, per una delle leggi di assorbimento,  $1 = 1 \vee (1 \wedge a) = 1 \vee a$ ; analogamente si procede per 0.

#### 4. ANELLI BOOLEANI E ALGEBRE DI BOOLE

In questa sezione arriveremo a provare che le nozioni di anello booleano e di reticolo booleano (ovvero di algebra di Boole) sono in sostanza interscambiabili, nel senso che si può costruire una struttura di reticolo booleano su ogni anello booleano e, viceversa, una struttura di anello booleano su ogni reticolo booleano, in modo che queste due costruzioni siano l'una inversa dell'altra.

In primo luogo, partendo da un anello booleano  $(R, +, \cdot)$  vogliamo definire una struttura di reticolo booleano su  $R$ . L'esempio dell'anello delle parti di un insieme può suggerirci in che modo procedere. Fissato un insieme  $S$ , infatti,  $(\mathcal{P}(S), \Delta, \cap)$  è un anello booleano ma  $\mathcal{P}(S)$  è anche un reticolo booleano, con operazioni reticolari  $\cup$  e  $\cap$ . La seconda operazione reticolare è proprio l'operazione di moltiplicazione nell'anello. Anche la prima operazione reticolare si può esprimere in termini delle operazioni dell'anello: per ogni  $A, B \in \mathcal{P}(S)$  abbiamo infatti  $A \cup B = (A \Delta B) \cup (A \cap B) = (A \Delta B) \Delta (A \cap B)$ . Inoltre il minimo ed il massimo del reticolo sono  $\emptyset$  e  $S$ , cioè lo zero e l'unità dell'anello, e ciascun  $A \in \mathcal{P}(S)$  ha come complemento, nel reticolo  $(\mathcal{P}(S), \subseteq)$ , l'insieme  $S \setminus A = S \Delta A = 1_{\mathcal{P}(S)} \Delta A$ .

Passando ora ad un arbitrario anello booleano  $(R, +, \cdot, 0_R, 1_R)$ , dove  $0_R$  e  $1_R$  sono lo zero e l'unità dell'anello, l'esempio di  $\mathcal{P}(S)$  suggerisce di definire in  $R$  l'operazione binaria  $\vee$  ponendo, per ogni  $a, b \in R$ ,

$$a \vee b := a + b + ab$$

e l'applicazione  $'$ :  $a \in R \mapsto 1_R + a \in R$  da utilizzare come operazione unaria di complemento.

**Proposizione 14.** *Con le notazioni appena fissate,  $(R, \vee, \cdot, 0_R, 1_R, ')$  è un'algebra di Boole.*

*Dimostrazione.* Dobbiamo verificare che  $(R, \vee, 0_R)$  e  $(R, \cdot, 1_R)$  siano monoidi commutativi, che valgano per  $\vee$  e  $\cdot$  le leggi di assorbimento e le proprietà distributive,<sup>(6)</sup> ed infine che l'applicazione  $'$  verifichi la condizione richiesta dalla definizione di complemento.

Che  $\vee$  sia commutativa è evidente, ed è anche chiaro che  $a \vee 0_R = a + 0_R + a0_R = a$  per ogni  $a \in R$ , quindi  $0_R$  è neutro rispetto a  $\vee$ . Proviamo l'associatività di  $\vee$ : per ogni  $a, b, c \in R$  si ha  $(a \vee b) \vee c = (a + b + ab) \vee c = (a + b + ab) + c + (a + b + ab)c = a + b + c + ab + ac + bc + abc$ . Si ha quindi  $a \vee (b \vee c) = (b \vee c) \vee a = b + c + a + bc + ba + ca + bca$ ; dunque  $(a \vee b) \vee c = a \vee (b \vee c)$ . È così provato che  $\vee$  è associativa;  $(R, \vee, 0_R)$  è quindi un monoide commutativo. Che lo sia anche  $(R, \cdot, 1_R)$  è già noto in partenza, dal momento che  $R$  è un anello booleano.

Verifichiamo le leggi di assorbimento. Per ogni  $a, b \in R$ ,  $a \vee (ab) = a + ab + a(ab)$ . Dal momento che  $R$  è booleano,  $a(ab) = a^2b = ab$  e  $ab + ab = 0_R$ , quindi  $a \vee (ab) = a + ab + ab = a$ . Inoltre  $a(a \vee b) = a(a + b + ab) = a^2 + ab + a^2b = a + ab + ab = a$ . Le leggi di assorbimento sono così provate. A questo punto già possiamo concludere che  $(R, \vee, \cdot)$  è un reticolo limitato.

Verifichiamo ora che  $\cdot$  è distributiva rispetto a  $\vee$ . Per ogni  $a, b, c \in R$  si ha  $a(b \vee c) = a(b + c + bc) = ab + ac + abc$  e  $(ab) \vee (ac) = ab + ac + (ab)(ac) = ab + ac + abc$ , dunque  $a(b \vee c) = (ab) \vee (ac)$ . Pertanto, utilizzando anche la proprietà commutativa, possiamo concludere che  $\cdot$  è distributiva rispetto a  $\vee$ .

Anche se non è strettamente necessario, verifichiamo anche che  $\vee$  è distributiva rispetto a  $\cdot$ . Per ogni  $a, b, c \in R$  abbiamo  $a \vee (bc) = a + bc + abc$  e  $(a \vee b)(a \vee c) = (a + b + ab)(a + c + ac) = a + ac + ac + ab + bc + abc + ab + abc + abc = a + bc + abc = a \vee (bc)$ . Dunque,  $\vee$  è distributiva rispetto a  $\cdot$ .

Resta infine da dimostrare che, per ogni  $a \in R$ , l'immagine di  $a$  mediante l'applicazione  $'$ , vale a dire  $a' := 1_R + a$ , verifica le condizioni  $a \vee (1_R + a) = 1_R$  e  $aa' = 0_R$ . Questo è molto facile: per ogni  $a \in R$  si ha  $aa' = a(1_R + a) = a + a = 0_R$  e  $a \vee a' = a + a' + aa' = a + (1_R + a) + 0_R = 1_R$ , come richiesto. Con questo la dimostrazione è completa  $\square$

Descriviamo ora la costruzione inversa: quella di un anello booleano a partire da un'algebra di Boole. Anche in questo caso ci facciamo guidare dall'esempio dell'algebra  $(\mathcal{P}(S), \cup, \cap, \emptyset, S, ^c)$  delle parti di un insieme  $S$  (come nell'[Esempio 10](#), il simbolo  $^c$  rappresenta l'operazione unaria di complemento in  $S$ ). Delle due operazioni binarie dell'anello (booleano)  $(\mathcal{P}(S), \Delta, \cap)$ , quella di moltiplicazione,  $\cap$ , è già tra le operazioni dell'algebra di Boole. Per esprimere l'altra, la differenza simmetrica, utilizzando le operazioni dell'algebra di Boole ci è utile osservare che se  $A$  e  $B$  sono parti di  $S$ , allora  $A \setminus B = A \cap (S \setminus B) = A \cap B^c$ . Dunque  $A \Delta B$  può essere scritta come  $(A \setminus B) \cup (B \setminus A) = (A \cap B^c) \cup (B \cap A^c)$  o anche come  $(A \cup B) \setminus (A \cap B) = (A \cup B) \cap (A \cap B)^c$ . Queste esempi suggeriscono due possibili modi per definire, in un'arbitraria algebra di Boole  $(L, \vee, \wedge, 0, 1, ')$ , un'operazione binaria di addizione + analoga alla differenza simmetrica in  $\mathcal{P}(S)$ . Il prossimo lemma mostra che queste due possibilità portano allo stesso risultato.

**Lemma 15.** *Sia  $(L, \vee, \wedge, 0, 1, ')$  un'algebra di Boole. Allora, per ogni  $a, b \in L$  si ha  $(a \wedge b') \vee (a' \wedge b) = (a \vee b) \wedge (a \wedge b)'$ .*

*Dimostrazione.* Usando la proprietà distributiva di  $\vee$  rispetto a  $\wedge$ , abbiamo:

$$(a \wedge b') \vee (a' \wedge b) = (a \vee a') \wedge (a \vee b) \wedge (b' \vee a') \wedge (b' \vee b) = 1 \wedge (a \vee b) \wedge (b' \vee a') \wedge 1 = (a \vee b) \wedge (a' \vee b') = (a \vee b) \wedge (a \wedge b)',$$

avendo utilizzato, per l'ultimo passaggio, una delle leggi di De Morgan ([Proposizione 12](#) (v)).  $\square$

---

<sup>(6)</sup>in realtà, per un'osservazione fatta a margine della definizione di reticolo distributivo, basterebbe dimostrare una sola delle due proprietà distributive.

Anche quest'altra osservazione può essere utile:

 **Lemma 16.** Sia  $(L, \vee, \wedge, 0, 1, {}')$  un'algebra di Boole. Allora, per ogni  $a, b \in L$  si ha  $(a \wedge b)' \wedge (a \wedge c) = a \wedge b' \wedge c$ .

*Dimostrazione.* Utilizzando una delle formule di De Morgan,  $(a \wedge b)' \wedge (a \wedge c) = (a' \vee b') \wedge a \wedge c = ((a' \wedge a) \vee (b' \wedge a)) \wedge c = (0 \vee (b' \wedge a)) \wedge c = a \wedge b' \wedge c$ .  $\square$

**Proposizione 17.** Sia  $(L, \vee, \wedge, 0, 1, {}')$  un'algebra di Boole. Se  $+$  è l'operazione binaria definita in  $L$  ponendo, per ogni  $a, b \in L$ ,  $a + b = (a \wedge b') \vee (a' \wedge b)$ , allora  $(L, +, \wedge)$  è un anello booleano, con zero 0 e unità 1.

 *Dimostrazione.* Iniziamo col verificare che  $(L, +)$  è un gruppo abeliano. Poiché  $\vee$  e  $\wedge$  sono commutative, è evidente che  $+$  è commutativa. Per ogni  $a, b, c \in L$  abbiamo:

$$\begin{aligned} (a + b) + c &= ((a \wedge b') \vee (a' \wedge b)) + c \\ &= (((a \wedge b') \vee (a' \wedge b)) \wedge c') \vee (((a \wedge b') \vee (a' \wedge b))' \wedge c) && [\text{usando il Lemma 15}] \\ &= (((a \wedge b') \vee (a' \wedge b)) \wedge c') \vee (((a \vee b) \wedge (a \wedge b)')' \wedge c) \\ &= (((a \wedge b') \vee (a' \wedge b)) \wedge c') \vee (((a' \wedge b') \vee (a \wedge b)) \wedge c) && [\text{usando la Proposizione 12}] \\ &= ((a \wedge b' \wedge c') \vee (a' \wedge b \wedge c')) \vee ((a' \wedge b' \wedge c) \vee (a \wedge b \wedge c)) \\ &= (a \wedge b' \wedge c') \vee (a' \wedge b \wedge c') \vee (a' \wedge b' \wedge c) \vee (a \wedge b \wedge c). \end{aligned}$$

Poiché  $+$  è commutativa, abbiamo quindi anche  $a + (b + c) = (b + c) + a = (b \wedge c' \wedge a') \vee (b' \wedge c \wedge a') \vee (b' \wedge c' \wedge a) \vee (b \wedge c \wedge a)$  ed allora, per la commutatività di  $\wedge$  e  $\vee$ ,  $a + (b + c) = (a + b) + c$ . Pertanto  $+$  è associativa. Per ogni  $a \in L$  vale  $a + 0 = (a \wedge 0') \vee (a' \wedge 0) = (a \wedge 1) \vee 0 = a \vee 0 = a$ , quindi 0 è neutro rispetto a  $+$ . Inoltre, sempre per ogni  $a \in L$ ,  $a + a = (a \wedge a') \vee (a' \wedge a) = 0 \vee 0 = 0$ , quindi ogni elemento di  $L$ , rispetto a  $+$ , è il simmetrico di sé stesso. Dunque  $(L, +)$  è un gruppo abeliano, con elemento neutro 0.

Sappiamo dalla definizione di algebra di Boole che  $(L, \wedge, 1)$  è un monoide commutativo. Verifichiamo la distributività di  $\wedge$  rispetto a  $+$ . Per ogni  $a, b, c \in L$  abbiamo:

$$a \wedge (b + c) = a \wedge ((b \wedge c') \vee (b' \wedge c)) = (a \wedge b \wedge c') \vee (a \wedge b' \wedge c)$$

e, utilizzando due volte il Lemma 16,

$$(a \wedge b) + (a \wedge c) = ((a \wedge b) \wedge (a \wedge c)') \vee ((a \wedge b)' \wedge (a \wedge c)) = (a \wedge b \wedge c') \vee (a \wedge b' \wedge c),$$

quindi  $a \wedge (b + c) = (a \wedge b) + (a \wedge c)$ .

A questo punto abbiamo dimostrato che  $(L, +, \wedge)$  è un anello (commutativo) unitario, con 0 e 1 come zero e unità. Per ogni  $a \in L$  vale  $a^2 = a \wedge a = a$ , quindi questo anello è booleano. La dimostrazione è così completa.  $\square$

Abbiamo così visto che ogni anello booleano  $(R, +, \cdot)$  determina una struttura di algebra di Boole sul suo stesso sostegno:  $(R, \vee, \cdot, 0_R, 1_R, {}')$ , definita come nella Proposizione 14. Per la Proposizione 17, questa definisce a sua volta un anello booleano, indichiamolo come  $(R, \oplus, \cdot)$ , con operazione additiva  $\oplus$  definita da: per ogni  $a, b \in R$

$$a \oplus b = (ab') \vee (a'b) \in R.$$

Ora, scelti comunque  $a, b \in R$ , poiché, in accordo con la Proposizione 14, per ogni  $u, v \in R$ , abbiamo  $u' = 1_R + u$  (e  $uu' = 0_R$ ) e  $u \vee v = u + v + uv$ ,

$$(ab') \vee (a'b) = (ab') + (a'b) + (ab')(a'b) = (ab') + (a'b) + aa'bb' = a(1_R + b) + (1_R + a)b + 0_R = a + ab + b + ab = a + b,$$

ricordando che  $ab + ab = 0_R$ . Dunque, l'operazione additiva  $\oplus$  dell'anello booleano costruito a partire da  $(R, \vee, \cdot, 0_R, 1_R, {}')$  non è altro che l'originale addizione in  $(R, +, \cdot)$ .

Questo significa che, dato un anello booleano  $R$ , se si costruisce un'algebra di Boole su  $R$  come indicato nella Proposizione 14 e poi, a partire da quest'ultima, si costruisce un anello booleano come indicato nella Proposizione 17, questo anello è precisamente l'anello  $R$  da cui si era partiti.

Lo stesso vale se si fa il discorso inverso. Se, partendo da un'algebra di Boole  $(L, \vee, \wedge, 0, 1, {}')$ , si definisce l'anello booleano  $(L, +, \wedge)$  come nella Proposizione 17 e poi si usa la Proposizione 14 per costruire un'algebra di Boole  $(L, \vee, \wedge, 0, 1, {}'')$  a partire da questo anello, l'algebra così ottenuta è quella originaria. Per provarlo, basta verificare che l'operazione  $\vee$  coincide con  $\vee$ . Infatti, una volta stabilito ciò, si ha che le due strutture di reticolo booleano su  $L$ , l'originale  $(L, \vee, \wedge)$  e la "nuova"  $(L, \vee, \wedge)$ , coincidono, quindi lo stesso è vero per le corrispondenti algebre di Boole.

 Verifichiamo  $\gamma = \vee$ . Scelti comunque  $a, b \in L$ , le costruzioni in [Proposizione 17](#) e [Proposizione 14](#) danno  $a + b = (a \wedge b') \vee (a' \wedge b)$  e  $a \gamma b = a + b + (a \wedge b)$ . Ma  $b = 1 \wedge b$  e, usando la proprietà distributiva di  $\wedge$  rispetto a  $+$ ,

$$\begin{aligned} a \gamma b &= a + (1 \wedge b) + (a \wedge b) = a + ((1 + a) \wedge b) \\ &= a + (a' \wedge b) && [\text{perché } a' = 1 + a] \\ &= (a \wedge (a' \wedge b')) \vee (a' \wedge a' \wedge b) \\ &= (a \wedge (a \vee b')) \vee (a' \wedge b) && [\text{legge di De Morgan}] \\ &= a \vee (a' \wedge b) && [\text{legge di assorbimento}] \\ &= (a \vee a') \wedge (a \vee b) = 1 \wedge (a \vee b) = a \vee b. \end{aligned}$$

Quindi, effettivamente,  $\gamma$  coincide con  $\vee$ . Possiamo sintetizzare quanto abbiamo dimostrato nel seguente teorema.

**Teorema 18.** Sia  $L$  un insieme. Sia  $\mathcal{A}$  l'insieme delle coppie ordinate  $(\vee, \wedge)$  di operazioni binarie in  $L$  che strutturano  $L$  come algebra di Boole, e sia  $\mathcal{B}$  l'insieme delle coppie ordinate  $(+, \cdot)$  di operazioni binarie in  $L$  che strutturano  $L$  come anello booleano. Allora le costruzioni descritte dalla [Proposizione 14](#) e dalla [Proposizione 17](#) definiscono due applicazioni, da  $\mathcal{B}$  a  $\mathcal{A}$  e da  $\mathcal{A}$  a  $\mathcal{B}$ , che sono l'una inversa dell'altra, quindi biettive.

Questa corrispondenza tra algebre di Boole e anelli booleani conserva la nozione di isomorfismo.

**Proposizione 19.** Siano  $(L_1, \vee_1, \wedge_1, 0_1, 1_1, ')$  e  $(L_2, \vee_2, \wedge_2, 0_2, 1_2, '')$  algebre di Boole e  $(L_1, +_1, \wedge_1)$  e  $(L_2, +_2, \wedge_2)$  i corrispondenti (nel senso del [Teorema 18](#)) anelli booleani. Sia poi  $f: L_1 \rightarrow L_2$  un'applicazione biettiva. Allora  $f$  è un isomorfismo di algebre di Boole se e solo se è un isomorfismo di anelli booleani.

*Dimostrazione.* Sia  $f$  un isomorfismo di algebre di Boole. Allora, per ogni  $a, b \in L_1$  si ha

$$f(a +_1 b) = f((a \vee_1 b') \wedge_1 (a' \vee_1 b)) = (f(a) \vee_2 f(b)'') \wedge_2 (f(a)' \vee_2 f(b)) = f(a) +_2 f(b)$$

e, ovviamente,  $f(a \wedge_1 b) = f(a) \wedge_2 f(b)$ . Quindi  $f$  è un isomorfismo di anelli booleani. Viceversa, se  $f$  è un isomorfismo di anelli booleani, allora, per ogni  $a, b \in L_1$  si ha  $f(a \wedge_1 b) = f(a) \wedge_2 f(b)$  e

$$f(a \vee_1 b) = f(a +_1 b +_1 (a \wedge_1 b)) = f(a) +_2 f(b) +_2 (f(a) \wedge_2 f(b)) = f(a) \vee_2 f(b).$$

Dunque  $f$  conserva le operazioni reticolari ed è quindi un isomorfismo di reticolati da  $L_1$  a  $L_2$ . Come già sappiamo dalla [sezione 4](#), da ciò segue che  $f$  è un isomorfismo di algebre di Boole.  $\square$

A questo punto possiamo concludere che lo studio degli anelli booleani equivale a quello delle algebre di Boole, e quindi a quello dei reticolati booleani. Vediamo anche che le sottoalgebre di Boole corrispondono precisamente ai sottoanelli unitari.

**Proposizione 20.** Sia  $(L, \vee, \wedge, 0, 1, ')$  un'algebra di Boole e sia  $(L, +, \wedge)$  il corrispondente anello booleano (nel senso del [Teorema 18](#)). Sia  $K \subseteq L$ . Allora  $K$  è una sottoalgebra di Boole di  $(L, \vee, \wedge, 0, 1, ')$  se e solo se è un sottoanello unitario di  $(L, +, \wedge)$ .

*Dimostrazione.* Sia  $K$  una sottoalgebra di Boole. Allora, per ogni  $a, b \in K$ , si ha  $a + b = (a \wedge b') \vee (a' \wedge b) \in K$ , quindi  $K$  è chiusa rispetto a  $+$ . Dal momento che  $(L, +, \wedge)$  è booleano, ogni elemento di  $L$  coincide col suo opposto. Da ciò segue che  $K$  è un sottogruppo di  $(L, +)$ . Ovviamente, poiché  $K$  è una sottoalgebra,  $K$  è un sottomonoide di  $(K, \wedge, 1)$ . Dunque,  $K$  è un sottoanello unitario di  $(L, +, \wedge)$ .

Viceversa, se  $K$  è un sottoanello unitario di  $(L, +, \wedge)$ , allora  $K$  è un sottomonoide di  $(L, \wedge, 1)$ . Inoltre  $0 \in K$ , per ogni  $a \in K$  si ha  $a' = 1 + a \in K$  e dunque, per ogni  $a, b \in K$ ,  $a \vee b = (a \wedge b') \vee (a' \wedge b) \in K$ . Concludiamo che  $K$  è anche un sottomonoide di  $(L, \vee, 0)$  e contiene il complemento in  $L$  di ogni suo elemento. Dunque,  $K$  è una sottoalgebra di Boole di  $(L, \vee, \wedge, 0, 1, ')$ .  $\square$

Da questi ultimi risultati e dal teorema di Stone per anelli booleani seguono subito i teoremi di Stone per algebre di Boole e per reticolati booleani.

**Teorema di Stone** (per algebre di Boole). Sia  $L$  un'algebra di Boole. Allora:

- (i) esiste un insieme  $S$  tale che  $L$  sia isomorfa ad una sottoalgebra di Boole dell'algebra delle parti  $(\mathcal{P}(S), \cup, \cap, \emptyset, S, {}^c)$  di  $S$ ;
- (ii) se  $L$  è finita, esiste un insieme  $S$  tale che  $L$  sia isomorfa all'algebra delle parti  $(\mathcal{P}(S), \cup, \cap, \emptyset, S, {}^c)$  di  $S$ .

**Teorema di Stone** (per reticolati booleani). Sia  $L$  un reticolo booleano. Allora:

- (i) esiste un insieme  $S$  tale che  $L$  sia isomorfo ad un sottoreticolo del reticolo  $(\mathcal{P}(S), \subseteq)$  delle parti di  $S$ ;
- (ii) se  $L$  è finito, esiste un insieme  $S$  tale che  $L$  sia isomorfo al reticolo  $(\mathcal{P}(S), \subseteq)$  delle parti di  $S$ .

Naturalmente, in conseguenza di questi teoremi, valgono anche per le algebre di Boole finite ed i reticolati booleani finiti le conseguenze osservate nel [Corollario 3](#) per gli anelli booleani: tutte le algebre di Boole finite e tutti i reticolati booleani finiti hanno per cardinalità una potenza di 2; se due algebre di Boole finite sono equipotenti (cioè hanno lo stesso numero di elementi) allora esse sono isomorfe; se due reticolati booleani finiti sono equipotenti allora essi sono isomorfi.

## 5. ANELLI BOOLEANI, STRINGHE DI ZERI E UNO ED OPERAZIONI BIT A BIT

In questa sezione faremo alcune osservazioni ed esempi su una delle situazioni in cui, in informatica, capita di incontrare strutture booleane.

Iniziamo da un accenno ad una costruzione generale. Fissiamo un anello  $R$  ed un intero positivo  $n$ . L'insieme  $R^n$  delle  $n$ -ple di elementi di  $R$  si può strutturare come anello definendo operazioni binarie di addizione e moltiplicazione “componente per componente”, cioè in questo modo: per ogni  $\underline{a} = (a_1, a_2, \dots, a_n)$  e  $\underline{b} = (b_1, b_2, \dots, b_n)$  appartenenti a  $R$  si pone

$$\underline{a} + \underline{b} = (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n) \quad \text{e} \quad \underline{a} \cdot \underline{b} = (a_1 b_1, a_2 b_2, \dots, a_n b_n); \quad (*)$$

abbiamo indicato con  $+$  e  $\cdot$  sia le operazioni in  $R$  che quelle in  $R^n$ . È un semplice [esercizio](#) la verifica del fatto che in questo modo  $R^n$  si struttura come anello (con  $(0_R, 0_R, \dots, 0_R)$  come zero e, se  $R$  è unitario,  $(1_R, 1_R, \dots, 1_R)$  come unità) e che  $R^n$  è booleano se  $R$  è booleano.

Consideriamo il caso in cui  $R$  è l'anello  $\mathbb{Z}_2$  degli interi modulo 2. Dal momento che  $\mathbb{Z}_2$  è booleano, anche  $\mathbb{Z}_2^n$  è booleano. I suoi elementi sono le  $n$ -ple di elementi di  $\mathbb{Z}_2$ , quindi le  $n$ -ple  $(a_1, a_2, \dots, a_n)$  dove ciascuno degli  $a_i$  è uno dei due elementi di  $\mathbb{Z}_2$ : o  $[0]_2$  oppure  $[1]_2$ . Ovviamente  $|\mathbb{Z}_2^n| = 2^n$ . Per semplificare la notazione possiamo scrivere 0 e 1 per  $[0]_2$  e  $[1]_2$  e rappresentare le  $n$ -ple come stringhe di lunghezza  $n$ , vale a dire, se, ad esempio,  $n = 5$ , scriviamo ‘ $a_1 a_2 a_3 a_4 a_5$ ’ piuttosto che  $(a_1, a_2, a_3, a_4, a_5)$ . Per esempio, sempre per  $n = 5$ , la stringa ‘10100’ sta per  $([1]_2, [0]_2, [1]_2, [0]_2, [0]_2) \in \mathbb{Z}_2^5$ .

Con queste notazioni, dette  $\underline{a} = 'a_1 a_2 \dots a_n'$  e  $\underline{b} = 'b_1 b_2 \dots b_n'$  due stringhe appartenenti a  $\mathbb{Z}_2^n$ , abbiamo  $\underline{a} + \underline{b} = 's_1 s_2 \dots s_n'$  e  $\underline{a} \cdot \underline{b} = 'p_1 p_2 \dots p_n'$ , dove, per ogni  $i \in \{1, 2, \dots, n\}$ ,  $s_i$  è la somma e  $p_i$  il prodotto di  $a_i$  e  $b_i$  in  $\mathbb{Z}_2$ , quindi  $s_i = 0$  se  $a_i = b_i$  e  $s_i = 1$  se  $a_i \neq b_i$ , mentre  $p_i = 1$  se  $a_i = b_i = 1$  e  $p_i = 0$  negli altri casi.

Molto probabilmente, chi legge riconosce in queste regole di calcolo le operazioni ‘bit a bit’ su “stringhe di zeri e uno” di fissata lunghezza con cui funzionano gli elaboratori elettronici, associate ai connettivi (operatori) logici XOR e AND. Quello che stiamo qui dicendo è che

*queste operazioni ‘bit a bit’ non sono altro che le due operazioni binarie dell’anello booleano  $\mathbb{Z}_2^n$ .*

Naturalmente lo stesso discorso si può estendere alle operazioni che, ai sensi della [Proposizione 14](#), strutturano  $\mathbb{Z}_2^n$  come algebra di Boole. Indicando con  $\underline{1}$  la stringa ‘11…1’ (di lunghezza  $n$ ), che è l’unità di  $\mathbb{Z}_2^n$ , il complemento di  $\underline{a} = 'a_1 a_2 \dots a_n'$  in  $\mathbb{Z}_2^n$  sarà  $\underline{1} + \underline{a}$  che, come si verifica subito, è la stringa ottenuta sostituendo in  $\underline{a}$  ogni 0 con 1 ed ogni 1 con 0; quello che abbiamo descritto è l’operatore NOT. Se poi  $\underline{b} = 'b_1 b_2 \dots b_n' \in \mathbb{Z}_2^n$ , è facile verificare che  $\underline{a} \vee \underline{b} = 'l_1 l_2 \dots l_n'$ , dove, per ogni  $i$ ,  $l_i = 0$  se  $a_i = b_i = 0$  e  $l_i = 1$  altrimenti; un modo per farlo è osservare che per le leggi di De Morgan ([Proposizione 12](#))  $\underline{a} \vee \underline{b} = \underline{1} + ((\underline{1} + \underline{a})(\underline{1} + \underline{b}))$ . Dunque  $\vee$  coincide con l’operatore OR bit a bit.

**Esercizio 21.** Verificare la correttezza della definizione dell’anello  $R^n$  data all’inizio di questa sezione e le proprietà di  $R^n$  li indicate.

Più in generale, siano  $n$  un intero positivo e  $(R_1, +_1, \cdot_1), (R_2, +_2, \cdot_2), \dots, (R_n, +_n, \cdot_n)$  anelli. Sia  $P$  il prodotto cartesiano  $R_1 \times R_2 \times \dots \times R_n$  e definiamo in  $P$  due operazioni binarie  $+$  e  $\cdot$  ponendo, per ogni  $\underline{a} = (a_1, a_2, \dots, a_n), \underline{b} = (b_1, b_2, \dots, b_n) \in P$ , come in  $(*)$ ,  $\underline{a} + \underline{b} = (a_1 +_1 b_1, a_2 +_2 b_2, \dots, a_n +_n b_n)$  e  $\underline{a} \cdot \underline{b} = (a_1 \cdot_1 b_1, a_2 \cdot_2 b_2, \dots, a_n \cdot_n b_n)$ . Verificare che  $(P, +, \cdot)$  è un anello e che questo anello è unitario (rispettivamente, commutativo, booleano) se ciascuno degli anelli  $R_i$  ha la stessa proprietà.

Infine, vediamo cosa di altro possiamo dire sull’anello  $\mathbb{Z}_2^n$  alla luce del teorema di Stone. Innanzitutto, (continuando a indicare con  $n$  un intero positivo fissato) cosa sono le  $n$ -ple di elementi di  $\mathbb{Z}_2$ ? Una maniera per rispondere è assumere che, per definizione, una  $n$ -pla di elementi di  $\mathbb{Z}_2$  sia un’applicazione da  $S := \{1, 2, \dots, n\}$  a  $\mathbb{Z}_2$ : la  $n$ -pla  $\underline{a} = 'a_1 a_2 \dots a_n'$  è l’applicazione  $i \in S \mapsto a_i \in \mathbb{Z}_2$ . Quindi l’insieme  $\mathbb{Z}_2^n$  è l’insieme  $\mathbb{Z}_2^S$  delle applicazioni da  $S$  a  $\mathbb{Z}_2$ .

A questo punto ci viene in aiuto la nozione insiemistica di funzione caratteristica. Se  $X$  è una parte di  $S$ , la funzione caratteristica di  $X$  in  $S$  (a valori in  $\mathbb{Z}_2$ ) è l’applicazione

$$\chi_{X,S}: i \in S \longmapsto \begin{cases} 1, & \text{se } i \in X \\ 0, & \text{se } i \notin X \end{cases} \in \mathbb{Z}_2.$$

Si ricorda che l’applicazione  $X \in \mathcal{P}(S) \mapsto \chi_{X,S} \in \mathbb{Z}_2^S$  è biettiva; l’applicazione inversa è quella che associa, ad ogni applicazione  $f: S \rightarrow \mathbb{Z}_2$ , l’antiimmagine di  $\{1\}$  mediante  $f$ . Dunque, ricordano anche che  $\mathbb{Z}_2^S = \mathbb{Z}_2^n$  e continuando a scrivere gli elementi di questo insieme come stringhe di lunghezza  $n$ , la stringa (funzione) caratteristica di una parte di  $S$  è la stringa ‘ $a_1 a_2 \dots a_n$ ’, dove per ogni  $i \in S$  si ha  $a_i = 1$  se  $i \in X$  e  $a_i = 0$  se  $i \notin X$ ; viceversa, una stringa ‘ $a_1 a_2 \dots a_n$ ’ corrisponde all’insieme degli  $i \in S$  tali che  $a_i = 1$ .

Facciamo un esempio per chiarire ulteriormente questa coppia di applicazioni biettive. Assumendo  $n = 7$ ,

la stringa	la 7-pla	l’insieme
è	e corrisponde a	
‘1011010’	$([1]_2, [0]_2, [1]_2, [1]_2, [0]_2, [1]_2, [0]_2)$	$\{1, 3, 4, 6\}$

Ora, per il teorema di Stone l’anello  $(\mathbb{Z}_2^n, +, \cdot)$  è isomorfo all’anello delle parti di un insieme. Poiché  $|\mathbb{Z}_2^n| = 2^n$  e  $|S| = n$ , dobbiamo avere  $\mathbb{Z}_2^n \simeq (\mathcal{P}(S), \Delta, \cap)$ . In effetti, possiamo verificare che l’applicazione biettiva appena descritta è un isomorfismo.

**Proposizione 22.** Siano  $n$  un intero positivo e  $S = \{1, 2, \dots, n\}$ . L'applicazione  $\varphi$  che ad ogni parte  $X$  di  $S$  associa la stringa che rappresenta la funzione caratteristica di  $X$  in  $S$  è un isomorfismo di anelli booleani da  $(\mathcal{P}(S), \Delta, \cap)$  a  $(\mathbb{Z}_2^n, +, \cdot)$ .

**Dimostrazione.** Siano  $A$  e  $B$  parti di  $S$ , e siano  $\underline{a} = 'a_1 a_2 \dots a_n' = \varphi(A)$  e  $\underline{b} = 'b_1 b_2 \dots b_n' = \varphi(B)$ . Allora, per ogni  $i \in S$ ,  $a_i = 1$  se e solo se  $i \in A$  (risultando  $a_i = 0$  altrimenti); similmente  $b_i = 1$  se e solo se  $i \in B$ . Sia  $\underline{c} = 'c_1 c_2 \dots c_n' = \varphi(A \cap B)$ . Allora, per ogni  $i$ ,  $c_i = 1$  se e solo se  $i \in A \cap B$ , cioè se e solo se  $a_i = b_i = 1$ . Da ciò è chiaro che  $\underline{c} = \underline{a} \cdot \underline{b}$ . Sia poi  $\underline{d} = 'd_1 d_2 \dots d_n' = \varphi(A \Delta B)$ . Allora, per ogni  $i \in S$ ,  $d_i = 1$  se e solo se  $i \in A \Delta B$ , cioè se e solo se vale esattamente una tra  $i \in A$  e  $i \in B$ , cioè se e solo se, tra  $a_i$  e  $b_i$  uno è 1 e l'altro è 0. Se ne ricava:  $\underline{d} = \underline{a} + \underline{b}$ . Abbiamo così provato che  $\varphi$  è un isomorfismo.  $\square$

Possiamo in definitiva concludere che lavorare su stringhe di zeri e uno di fissata lunghezza  $n$  utilizzando le operazioni ‘bit a bit’ è del tutto equivalente a lavorare nell’anello (booleano) delle parti dell’insieme  $S$ . La moltiplicazione ‘bit a bit’ (operatore AND) corrisponde all’operazione di intersezione, l’addizione (modulo 2, operatore XOR) corrisponde all’operazione di differenza simmetrica.

Ad esempio, se, come sopra,  $n = 7$  e  $\underline{a} = '1011010'$ , ed inoltre  $\underline{b} = '0011100'$ , allora  $\underline{a}$  e  $\underline{b}$  corrispondono ai sottoinsiemi  $A = \{1, 3, 4, 6\}$  e  $B = \{3, 4, 5\}$  di  $\{1, 2, 3, 4, 5, 6, 7\}$ , e possiamo completare come segue una tabella in cui le stringhe al primo rigo corrispondono ad insiemi al secondo rigo:

$\underline{a} = '1011010'$	$\underline{b} = '0011100'$	$\underline{a} + \underline{b} = '1000110'$	$\underline{a} \cdot \underline{b} = '0011000'$
$A = \{1, 3, 4, 6\}$	$B = \{3, 4, 5\}$	$A \Delta B = \{1, 5, 6\}$	$A \cap B = \{3, 4\}$

# Strutture Booleane

Un reticolo si dice **BOOLEANO** se e solo se è complementato e distributivo.

Esempio:  $(P(S), \subseteq)$

Se  $(L, \leq)$  è un reticolo booleano, ogni elemento di  $L$  ha esattamente un complemento in esso.

In termini di operazioni reticolari (ed altre operazioni):

$$(L, \vee, \wedge, 0, 1, ')$$

min      max

$\vdash : a \in L \mapsto a' \in L$   
dove  $a'$  è il comp. direzionale di  $a$  in  $L$

1)  $(L, \vee, 0)$  e  $(L, \wedge, 1)$  sono monoidi commutativi.

2) leggi di assorbimento:  $\forall a, b, c \in L$   
 $a \wedge (a \vee b) = a = a \vee (a \wedge b)$

3) distributività di  $\vee$  risp.  $\wedge$   
e viceversa.

4)  $\forall a \in L$  ( $a \wedge a' = 0$ ,  $a \vee a' = 1$ ).

N.B. Il quale scambia  $\vee \leftrightarrow \wedge$  e  $0 \leftrightarrow 1$ , ma rimane un reticolo booleano.

## Regole di calcolo

$\forall a, b \in L$  (algebra di Bool)

- $a \wedge 0 = 0$
- $1' = 0$
- $a \vee 1 = 1$
- $(a')' = a$
- $(a \vee b)' = a' \wedge b'$
- $(a \wedge b)' = a' \vee b'$

De

Morgan

serve:

$$\begin{aligned} & (a \vee b) \vee (a' \wedge b') = 1 \quad \text{e } (a \vee b) \wedge (a' \wedge b') = 0 \\ & ((a \vee b) \vee a') \wedge ((a \vee b) \wedge b') = 1 \quad ((a \wedge (a' \wedge b')) \vee (b \wedge (a' \wedge b'))) = 0 \\ & ((a \vee a') \vee b) \wedge ((a \vee a') \wedge b) = 1 \quad = 0 \vee 0 = 0 \end{aligned}$$

# Anelli Booleani

$$(P(S), \subseteq)$$

$$(P(S), \cup, \cap, \emptyset, S, {}^c)$$

$${}^c : x \in P(S) \mapsto S \setminus x \in P(S)$$

$$(P(S), \Delta, \cap)$$

Un anello  $(R, +, -)$  è booleano se e solo se è unitario e ogni suo elemento è idempotente  $\Leftrightarrow (\forall a \in R \quad a^2 = a)$

Proprietà:

- 1) ogni anello booleano  $R$  è commutativo e verifica:  $\forall a \in R \quad (a^2 = O_R)$  (cioè  $\forall a \in R \quad (a = -a)$ )

DIM  $\forall a, b \in R$

$$(a+b)^2 = (a+b)(a+b) = a^2 + ab + ba + b^2 = a+a + ab + ba$$

$$ab + ba = O_R \Rightarrow ba = -(ab) = ab$$

Per  $b = a$ , otteniamo:

$$a^2 + a^2 = O_R$$

$$\therefore a^2 = a + a$$

$$a = -a$$



## Teorema di Stone

Sia  $R$  un anello booleano:

1) se  $R$  è finito, esiste  $S$  tale che  $R \cong (P(S), \Delta, \cap)$

2) in ogni caso, esiste  $S$  tale che  $R$  sia isomorfo ad un sottanello unitario di  $(P(S), \Delta, \cap)$

$\{ X \in P(\mathbb{N}) \mid X \text{ è finito oppure } \mathbb{N} \setminus X \text{ è finito} \}$  anello booleano non isomorfo  
a nessun  $P(S)$

## Corollario:

① Se  $R$  è un anello booleano finito,  $\exists n \in \mathbb{N} \quad (|R| = 2^n)$ .

② Se  $R$  e  $T$  sono anelli booleani finiti e quipotenti,  $R \cong T$ .

# Algebra di Boole

L'anello  $(R, +, \cdot)$  è booleano

Definiamo  $V: R \times R \rightarrow R$  via:  $\forall a, b \in R \quad (a \vee b := a+b+ab)$

- $V$  è commutativa
- $V$  è associativa
- ha  $0_R$  come neutro

$(R, V, 0_R)$  monoide commutativo

$(R, \cdot, 1_R)$  monoide commutativo

$$\forall a, b, c \in R \quad a \cdot b = ab = -(ab)$$

Leggi di assorbiamento:  $a \vee (ab) = a+ab+a^2b = a+ab+ab=a$

$$a \cdot (ab) = a(a+b+ab) = a^2+ab+a^2b=a$$

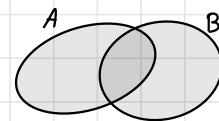
$$a \cdot (b \vee c) = a(b+c+bc) = ab+ac+abc$$

$$ab \vee ac = ab+ac+a^2bc = ab+ac+abc$$

$$\forall a \in R$$

$$a \cdot a' = a(1+a) = a+a^2 = a+a = 0_R$$

$$a \vee a' = a+a' + a \cdot a' = a+(1_R+a) + 0_R = 1_R$$



$$A \cup B = (A \Delta B) \Delta (A \cap B)$$

$$A^c = S \setminus A = S \Delta A = I_{P(S)} \Delta A$$

$$(A \cup B) \setminus (A \cap B)$$

$$A \cdot B = A \cap (S \setminus B)$$

$$A \Delta B = \langle$$

$$(A \setminus B) \cup (B \setminus A)$$

- $(L, V, \wedge, 0, 1, ')$  di Boole



$$\forall a, b \in L$$

$$a+b := (a \wedge b') \vee (a' \wedge b) = (a \vee b) \wedge (a \wedge b)' \quad \text{per De Morgan}$$

- $(L, +, \wedge)$  anello booleano

Sottanello unitario  $\leftrightarrow$  sottogruppo di Bool

chiuso risp.  $V \rightarrow 1 \leftrightarrow$  sottoreticolto

con 0 e 1 uguali

a cui appartenga il complem. di ogni elem.

# Algebra

## Lezione 28/11



## Principio d'induzione pt.2

$\forall X \in P(\mathbb{N}) \quad X \neq \emptyset \Rightarrow \exists \min(X, \leq)$  insieme ben ordinato

È la regola per cui vale il principio d'induzione

$$X = \{n \in \mathbb{N} \mid \neg p(n)\} \quad \begin{matrix} \text{insieme dei} \\ \leftarrow \text{controesempi} \end{matrix}$$

$$\forall n \in \mathbb{N} (p(n)) \xrightarrow{\text{negazione}} \exists m \in \mathbb{N} (\neg(p(m))), \text{ cioè } X \neq \emptyset, \text{ cioè } \exists \min(X, \leq).$$

### Induzione

$b \in \mathbb{N}$   $p$  predicato unario

$$N_b = \{x \in \mathbb{N} \mid b \leq x\}$$

$$(p(b) \wedge \forall n \in \mathbb{N} (p(n) \Rightarrow p(n+1))) \Rightarrow \forall n \in N_b (p(n))$$

DIM Sia  $X = \{n \in N_b \mid \neg(p(n))\}$

$$X \neq \emptyset \Rightarrow \exists m = \min(N_b, \leq).$$

Si ha  $m \neq b$ , perché abbiamo assunto  $p(b)$ , cioè  $b \notin X$ .

Allora  $m > b$  e così  $m-1 \in N_b$ . Ma  $m-1 < m = \min X$ ,

dunque  $m-1 \notin X$ : allora  $p(m-1)$  è vera. Per il passo induttivo,

ponendo  $m-1$  al posto di  $n$ , abbiamo  $p(m-1) \Rightarrow p((m-1)+1) = p(m)$

Verbo

Falso

Contraddizione! Dunque  $X = \emptyset$ , cioè  $\forall n \in N_b (p(n))$

### Seconda forma del principio di induzione

$$\forall t \in \mathbb{N}_b \quad M_t = \{x \in \mathbb{N}_b \mid x < t\} \quad \text{cioè: } M_t = [b, t]_{(\mathbb{N}, \leq)}$$

$$\left( \forall t \in \mathbb{N}_b \quad \left( \forall n \in M_t (p(n) \Rightarrow p(t)) \right) \right) \Rightarrow \forall n \in N_b (p(n))$$

## Elementi associati

Sia  $(M, \cdot, 1_M)$  monoidale commutativo

$$\forall a, b \in M \quad a|b \underset{(M, \cdot)}{\Leftrightarrow} \exists c \in M (b=ac) \quad a \sim b \Leftrightarrow a|b \wedge b|a \\ a \text{ e } b \text{ sono associati in } (M, \cdot)$$

$$\sim_{(M, \cdot)} \in Eq(M)$$

$$\forall a, a', b, b' \in M$$

Siano  $a|a'$  e  $b|b'$ . Allora  $a|b \Leftrightarrow a|b'$

DIM Se  $a|b$  abbiamo:  $a|a$ ,  $a|b$ ,  $b|b'$ , quindi  $a|b'$  per transitività.

In modo analogo,  $a|b' \Rightarrow a|a$ ,  $a|b'$ ,  $b'|b \Rightarrow a|b$ .



$$\forall a \in M$$

$$Div(a) = \{d \in M \mid d|a\} \quad a \cdot M = \{ax \mid x \in M\} = \{m \in M \mid a|m\}$$

Da ciò segue:

$$\forall a, b \in M \quad a \sim b \underset{(M, \cdot)}{\Leftrightarrow} Div(a) = Div(b) \Leftrightarrow a \cdot M = b \cdot M$$

DIM ( $\Rightarrow$ )  $\forall d \in Div(a)$

Se  $d|a$  e  $a|b$  segue  $d|b$   $\Rightarrow \forall d \in Div(b)$ ,  $d \in Div(a)$

( $\Leftarrow$ ) Poiché  $a \in Div(a)$ , se  $Div(b) = Div(a)$ , allora  $a \in Div(b)$ , cioè  $a|b$ .

Similmente  $b \in Div(b) = Div(a) \Rightarrow b|a$ , allora  $Div(a) = Div(b) \Rightarrow a \sim b$ .

## Alcune proprietà

•  $\forall u \in U(M) \quad \forall a \in M \quad (a \sim au)$  (infatti  $a|au$  e  $a = (au)u^{-1}$ , quindi  $au|a$ )

• Se  $a$  è cancellabile in  $M$ , si ha che

$$\{au \mid u \in U(M)\} = [a]_{\sim_{(M, \cdot)}} : l'insieme degli elementi di  $M$  associati ad  $a$ .$$

DIM Sappiamo già (1). Va ora verificato che ogni elemento associato ad  $a$  in  $M$  è della forma  $au$  per un opportuno  $u \in U(M)$ . Sia  $b \in M$ . Fissati tali  $u$  e  $v$

$$a|b \Rightarrow \exists u \in M (b = au) \quad a \cdot u = a = bv = (au)v = a(vu).$$

$$b|a \Rightarrow \exists v \in M (a = bv) \quad \text{Poiché } a \text{ è cancellabile, } 1_M = uv, \text{ quindi } v = u^{-1} \text{ e } u \in U(M).$$

## Monoidi cancellativi e divisori

Sia  $(R, +, \cdot)$  un dominio di integrità unitario.

$(R, \cdot)$  monoido commut.

$(R \setminus \{0_R\}, \cdot)$  monoide commut. cancellativo (o regolare): ogni elemento è cancellabile

$$\forall a \in R \quad [a]_{(R, \cdot)} = aU(R) = \{au \mid u \in U(R)\}$$

Sia  $a \in M$

Tra i divisori di  $a$  in  $M$  ci sono:

- 1) gli associati ad  $a$
- 2) gli elementi invertibili di  $M$

↑

$$\forall u \in U(M) \quad a = u(u^{-1}a) \Rightarrow u \mid a$$

DIVISORI  
BANALI DI  $a$

Un divisore di  $a$  si dice proprio,  
se esiste se non è associato ad  $a$ .

$a$  è irreducibile in  $(M, \cdot)$  se esiste se

- 1)  $a \in U(M)$
- 2)  $a$  non ha in  $M$  divisori non banali.

In  $\mathbb{Z}^*$ , corrispondono ai numeri primi.



## Fattorizzazioni

Scriveremo (supponendo si possa)  $a$  come prodotto di irreducibili:

$$a = p_1 p_2 \dots p_t \quad (\text{dove } t \in \mathbb{N}^* \text{ e } \forall i \in \{1, 2, \dots, t\} \quad p_i \text{ è irred. in } (M, \cdot))$$

$$a = q_1 q_2 \dots q_s \quad (\sim s \sim \dots \sim q_i \sim \dots)$$

Diciamo che queste due fattorizzazioni di  $a$  in prodotto di irreducibili sono essenzialmente uguali se e solo se:

1)  $s = t$

2) "i fattori  $p_i$  e  $q_i$  sono gli stessi a meno dell'ordine e della sostituzione di alcuni di essi con el. associati"

ovvero:

$$\text{In } (\mathbb{N}, \cdot) \quad 10 = 2 \cdot 5 = 5 \cdot 2$$

$\overset{\uparrow}{p_1} \quad \overset{\uparrow}{p_2} \quad \overset{\uparrow}{q_1} \quad \overset{\uparrow}{q_2}$

in  $(\mathbb{Z}, \cdot)$ , vale anche

$$10 = \underset{\sim p_1}{(-2)} \underset{\sim p_2}{(-5)}$$

vale a dire:  $\exists \sigma \in \text{Sym } \{1, 2, \dots, t\} \quad (\forall i \in \{1, 2, \dots, t\} \quad (q_{\sigma(i)} \sim p_i))$

## Monoide fattoriale

Un monoide commutativo  $(M, \cdot)$  si dice **fattoriale** se esiste:

- 1) è cancellativo
- 2) ogni elemento di  $M \setminus U(M)$  è prodotto di irriducibili in modo essenzialmente unico.  
cioè: due qualsiasi fattorizzazioni dell'elemento come prodotti di irriducibili sono essenzialmente uguali.

T. F. A. :  $(\mathbb{N}^*, \cdot, 1)$  è fattoriale

$(\mathbb{Z} \setminus \{0\}, \cdot, 1)$  è fattoriale

$(\mathbb{Z}, +, \cdot)$  è un anello fattoriale

Anello fattoriale

dominio di integrità unitario  $(R, +, \cdot)$

tal che  $(R \setminus \{0_R\}, \cdot)$  sia fattoriale.

# Algebra

## Lezione 30/11



MATH

## Teorema dei numeri primi

$$\forall n \in \mathbb{N}^* (n > 1 \Rightarrow n \text{ è prodotto di primi})$$

DIM Per contraddizione minima  $X = \{n \in \mathbb{N} \mid n > 1 \wedge n \text{ non è prodotto di primi}\}$

$X \neq \emptyset \Rightarrow \exists m = \min X$ . Allora  $m$  non è primo. Allora, perché  $m > 1$ ,  $m$  ha qualche divisore non banale  $a$ . Allora  $\exists b \in \mathbb{N}^* (m = ab)$ , inoltre  $1 < a < m$ , quindi anche  $b$  verifica sicuramente  $1 < b < m$ . Ne segue  $a, b \notin X$ , cioè  $a$  e  $b$  sono prodotti di primi. Contraddizione!!

Esercizio uguale per  $\mathbb{Z}$ :  $\forall n \in \mathbb{Z} \setminus \{0, 1, -1\} (n \text{ è prodotto di num. primi in } \mathbb{Z})$

$$\forall p \in \mathbb{Z} \quad p \text{ primo} \Rightarrow \forall a, b \in \mathbb{Z} \left( p \mid ab \Rightarrow (p \mid a \vee p \mid b) \right)$$

$$\begin{aligned} a &= p_1 \cdots p_n & \exists c \in \mathbb{Z} \quad (pc = ab) & c = r_1 \cdots r_t \\ b &= q_1 \cdots q_n & p_1 r_1 \cdots r_t = p_1 \cdots p_n q_1 \cdots q_n \end{aligned}$$

## Descrivere le fattorizzazioni

In  $\mathbb{N}^*$ , sia  $a = p_1^{\alpha_1} \cdots p_t^{\alpha_t}$  dove  $t \in \mathbb{N}^*$

$$\forall j, i \in \{1, 2, \dots, t\} \quad p_i \text{ è primo} \quad i \neq j \Rightarrow p_i \neq p_j \quad \alpha_i \in \mathbb{N}$$

I divisori di  $a$  in  $\mathbb{N}^*$  sono tutti esclusi i numeri della forma:

$$p_1^{\lambda_1} p_2^{\lambda_2} \cdots p_t^{\lambda_t}$$

$$\text{dove } \forall i \in \{1, 2, \dots, t\} \quad (\alpha_i \geq \lambda_i \in \mathbb{N})$$

Questi sono in tutto  $(\alpha_1 + 1)(\alpha_2 + 1) \cdots (\alpha_t + 1)$

$$12 = 2 \cdot 2 \cdot 3 = 2^2 \cdot 3^1 = 2^2 \cdot 3^1 \cdot 5^0$$

$$\text{Div}(12) = \{2^{\lambda_1} \cdot 3^{\lambda_2} \mid \lambda_1 \in \{0, 1, 2\} \wedge \lambda_2 \in \{0, 1\}\}$$

$$\lambda_1 = 0 \quad 1 \quad 3$$

$$\lambda_1 = 1 \quad 2 \quad 6$$

$$\lambda_1 = 2 \quad 4 \quad 12$$

Sia  $d \in \mathbb{N}^*$

$d \mid a \Rightarrow$  ogni primo che divide  $d$ , divide  $a$

Sia  $d = p_1^{\lambda_1} \cdots p_t^{\lambda_t} \mid a$ , possiamo porre  $a = db$  per un opportuno  $b \in \mathbb{N}^*$

Se  $\lambda_1 > \alpha_1$ ,  $p_1^{\lambda_1 - \alpha_1} p_2^{\lambda_2} \cdots p_t^{\lambda_t} b = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_t^{\alpha_t}$  assurdo! Dunque  $\lambda_1 \leq \alpha_1$

## Massimo Comun Divisore (M.C.D.) e Minimo Comune Multiplo (m.c.m.)

Siamo  $a, b \in \mathbb{N}^*$      $a = p_1^{\alpha_1} \cdots p_t^{\alpha_t}$      $b = p_1^{\beta_1} \cdots p_t^{\beta_t}$

I divisori comuni di  $a \cdot b$  sono i numeri della forma:

$$p_1^{\lambda_1} \cdots p_t^{\lambda_t}, \text{ dove } i \in \{1, 2, \dots, t\} \text{ e } \lambda_i \leq \min\{\alpha_i, \beta_i\} := S_x$$

Allora  $p_1^{\delta_1} p_2^{\delta_2} \cdots p_t^{\delta_t}$  è un MCD di  $\{a, b\}$

$(M, \cdot)$  monade commutativa.

$X \subseteq M$  dica un MCD( $X$ ) in  $(M, \cdot)$

Allora  $\forall h \in M$   $h$  è un MCD( $x$ )  $\Leftrightarrow h \sim d$  associati

In modo analogo  $p_1^{M_1} p_2^{M_2} \cdots p_t^{M_t}$  è un mcm  $\{a, b\}$

$$M_i = \max\{\alpha_i, \beta_i\}$$

$$S_x + M_x = \alpha_x + \beta_x$$

In  $\mathbb{Z}$   $ab = p_1^{\alpha_1 + \beta_1} p_2^{\alpha_2 + \beta_2} \cdots p_t^{\alpha_t + \beta_t} = dm$

In generale:  $ab \sim dm$

# Competibilità

$$\begin{array}{l} a \in (S, *) \mapsto [a]_\sigma \in (\mathbb{Z}_\sigma, \oplus) \text{ proiezione canonica} \\ (S, *) \quad \oplus: S \times S \rightarrow S \\ \sigma \in Eq(S) \quad \mathbb{Z}_\sigma \times \mathbb{Z}_\sigma \rightarrow \mathbb{Z}_\sigma \\ ([a]_\sigma, [b]_\sigma) \mapsto [a * b]_\sigma \end{array}$$

omomorfismo  
suriettivo

Esempio:

$$\begin{aligned} S &= \mathbb{Z} & S_\sigma &= \{\mathbb{N}^*, \{0\}, \mathbb{Z} \setminus \mathbb{N}\} & \xrightarrow{\sigma} \\ \mathbb{N}^* \oplus (\mathbb{Z} \setminus \mathbb{N}) &= [\mathbb{1} \setminus \mathbb{3}]_\sigma = [-2]_\sigma = \mathbb{Z} \setminus \mathbb{N} \\ [\mathbb{1}]_\sigma &= [\mathbb{-3}]_\sigma & &= [\mathbb{10} \setminus \mathbb{3}]_\sigma = [\mathbb{7}]_\sigma = \mathbb{N}^* \\ [\mathbb{10}]_\sigma & & & \text{NON VA BENE!} \end{aligned}$$

$\star$  è ben definita se e solo se  $\forall a, a', b, b' \in S$

$$([a]_\sigma, [b]_\sigma) = ([a']_\sigma, [b']_\sigma) \Rightarrow [a * b]_\sigma = [a' * b']_\sigma$$

ovvero:  $(a * a' \wedge b * b') \Rightarrow (a * b) \sigma (a' * b')$  oppure  $(a * b) \sigma (a' * b') \sigma (a' * b') \Rightarrow (a * b) \sigma (a' * b')$

Se questa condizione vale si dice che  $*$  e  $\sigma$  sono tra loro compatibili.

Esempio:

+ non è compatibile con  $\sigma$ , invece  $\cdot$  sì

$$a^c \sim b^c$$

$$a \equiv_z b \Rightarrow c \cdot a \sim c \cdot b$$

$\sigma$  è compatibile a sx con  $*$  (ovunque)  $\Leftrightarrow \forall a, b, c \in S (a * b \Rightarrow c * a \sim c * b)$

$\sigma$  è compatibile a dx con  $*$  (ovunque)  $\Leftrightarrow \forall a, b, c \in S (a * b \Rightarrow a * c \sim b * c)$



## Moduli e Congruenze

$\forall m \in \mathbb{Z}$  Definiamo:  $\equiv_m \in Rel(\mathbb{Z})$

$a \equiv_m b \pmod{m}$  significa  $a \equiv_m b$

$$\forall a, b \in \mathbb{Z} \quad a \equiv_m b \Leftrightarrow m | a - b$$

$\forall m \in \mathbb{Z}$ ,  $\equiv_m$  è:

- riflessiva:  $\forall a \in \mathbb{Z}$  in quanto, cioè  $a \equiv_m a$
- simmetrica:  $\forall a, b \in \mathbb{Z} m | a - b \Leftrightarrow m | -(a - b) = b - a$
- transitiva:  $\forall a, b, c \in \mathbb{Z} a \equiv_m b \wedge b \equiv_m c \Rightarrow a \equiv_m c$   
 $\Rightarrow m | a - b \wedge m | b - c$   
 $\text{dunque } m | (a - b) + (b - c) = a - c$  //

Parentesi:

Sia  $R$  un anello commutativo.

$$\forall a, b, c \in R \quad \forall u, v \in R$$

$$(a/b \wedge a/c) \Rightarrow a/bv + cv$$

DIM  $\exists \gamma, \beta \in R (a\beta = b \wedge a\gamma = c)$

$$a(bu + bv) = abu + avc = bu + cv$$

$\equiv_m \in Eq(\mathbb{Z})$  compatibile con +:

$$\forall a, b, c \in \mathbb{Z} \quad a \equiv_m b \Leftrightarrow a + c \equiv_m b + c$$

$$c+a \equiv_m c+b \Leftrightarrow m | (c+a) - (c+b) = a-b$$

$\equiv_m$  compatibile con  $\cdot$ :

$$\forall a, b, c \in \mathbb{Z} \quad a \equiv_m b \Leftrightarrow m | a \cdot b \Rightarrow m | c(a \cdot b) = ca \cdot cb \Rightarrow ca \equiv_m cb$$

$\equiv_m$  è una congruenza in  $(\mathbb{Z}, +, \cdot)$



$$\mathbb{Z}_m := \mathbb{Z}/\equiv_m \quad (\mathbb{Z}_m, +, \cdot) \text{ opp. indotte da quelle in } \mathbb{Z}$$

- $\equiv_0$  è la rel. di uguaglianza in  $\mathbb{Z}$
- $\equiv_1$  è la rel. universale in  $\mathbb{Z}$
- $\equiv_2$  è la rel. "avere la stessa parità"
- $\equiv_m = \equiv_{-m}$
- $\forall n \in \mathbb{Z} \quad \forall a, b \in \mathbb{Z} \quad (a \equiv_n b \Rightarrow \forall d \in \text{Div}_{(n)} \quad a \equiv_d b)$

$$\forall a, m \in \mathbb{Z}$$

$$[a]_m := [a]_{\equiv_m} = \{b \in \mathbb{Z} \mid a \equiv_m b\} =: a + m\mathbb{Z}$$

$$\forall b \in \mathbb{Z} \quad (b \equiv_m a \Leftrightarrow m | b-a \Leftrightarrow \exists k \in \mathbb{Z} \quad (mk = b-a) \Leftrightarrow \exists k \in \mathbb{Z} \quad (b = a + mk))$$

Sia  $m \in \mathbb{Z} \setminus \{0\}$   $\forall a \in \mathbb{Z} \quad [a]_m \cap \mathbb{N} \neq \emptyset$

$$\exists r = \min(\mathbb{N} \cap [a]_m) \quad a \bmod m := r$$

talè  $r$  verifica  $0 \leq r < |m|$

$$r \equiv_m a \quad r \equiv_m n - |m| \quad \text{quando} \quad n - |m| \in [a]_m$$

$n - |m| < r$ , dunque per definizione di  $r$   $n - |m| \notin \mathbb{N}$ ;  $n - |m| < 0$

Allora  $\forall a \in \mathbb{Z} \quad \exists r \in \mathbb{N} \quad (a \equiv_m r, r < |m|)$

$$\mathbb{Z}_m = \{[0]_m, [1]_m, [2]_m, \dots, [|m|-1]_m\}$$

$\text{Se } i, j \in \mathbb{N} \quad i < j < |m|$

$$m \nmid j-i \Rightarrow i \not\equiv_m j$$

# Algebra

## Lezione 02/12



N.B.

Queste pagine contengono, verso la fine, anche alcuni teoremi non enunciati in classe. Se si desiderano direttamente gli appunti relativi all'aritmetica modulare, basta saltare queste pagine. Se poi qualcosa dovesse non essere chiaro, qui potrebbe essere spiegato in maniera diversa e forse più efficace.

# Una introduzione all'aritmetica modulare

GIOVANNI CUTOLO

Questo articolo è pensato per studenti non universitari che abbiano qualche curiosità nei confronti di un argomento non abitualmente incontrato nei corsi scolastici. La trattazione è elementare e di tono informale quanto possibile, anche se non rinuncia del tutto a dare qualche indicazione di cosa si possa incontrare oltre questo primo approccio, e non ha alcuna pretesa di completezza. Sono presentate, nella seconda nella terza sezione dell'articolo, due applicazioni elementari dell'aritmetica modulare.

## 1. INTRODUZIONE INFORMALE

Siamo abituati a suddividere i numeri interi in due categorie: quella dei numeri pari e quella dei numeri dispari. Sappiamo anche, magari senza esserne del tutto consapevoli, che questa suddivisione ha una semplice ma importante proprietà di “buon comportamento” (*compatibilità* è la parola che si usa in matematica) nei confronti delle operazioni di addizione e moltiplicazione tra numeri interi. La proprietà è questa: dati due arbitrari numeri interi  $a$  e  $b$ , basta conoscere la parità di  $a$  e quella di  $b$  (cioè se  $a$  e  $b$  siano pari o dispari) per conoscere la parità di  $a + b$  e  $ab$ : se  $a$  e  $b$  sono entrambi pari o entrambi dispari allora  $a + b$  è pari, altrimenti  $a + b$  è dispari; se  $a$  e  $b$  sono entrambi dispari allora  $ab$  è dispari, altrimenti  $ab$  è pari. Questo fa sì che abbiano senso affermazioni come ‘pari più dispari fa dispari’, o ‘pari per dispari fa pari’, che tavolta si utilizzano.

Vale qualcosa del genere per altre “suddivisioni” (anche qui c’è un termine tecnico: nel linguaggio della teoria degli insiemi queste suddivisioni si chiamano *partizioni*) dell’insieme dei numeri interi? Vediamo: possiamo ripartire l’insieme degli interi in tre sottoinsiemi: quello dei numeri interi positivi, quello dei numeri interi negativi, quello che consiste del solo zero. Questa partizione è compatibile con la moltiplicazione: dati due numeri interi  $a$  e  $b$  per stabilire se  $ab$  è positivo, negativo o zero basta sapere se sono positivi, negativi o zero  $a$  e  $b$  (positivo per positivo dà positivo, negativo per zero dà zero etc.). Lo stesso non vale però per l’addizione: se, ad esempio,  $a$  è positivo e  $b$  è negativo, la loro somma  $a + b$  può essere negativa, zero o positiva ( $1 + (-2) < 0$ ;  $1 + (-1) = 0$ ;  $2 + (-1) > 0$ ). Diciamo quindi che la suddivisione degli interi tra positivi, negativi e zero non è compatibile con l’operazione di addizione tra interi.

Vediamo un altro esempio; questa volta, per semplicità, ci limiteremo agli interi positivi. Ripartiamo gli interi positivi per ultima cifra (quella delle unità; stiamo facendo riferimento alla consueta rappresentazione degli interi in base 10 che siamo abituati ad usare sin da piccoli). Supponiamo cioè di disporre gli interi positivi in dieci contenitori, ad esempio, dieci cassetti, ciascuno etichettato da una cifra  $(0, 1, 2, \dots, 9)$ , infilando nel cassetto etichettato da 0 tutti gli interi positivi con ultima cifra 0, in quello etichettato da 1 tutti gli interi positivi con ultima cifra 1 e così via: il cassetto con etichetta  $i$ , che possiamo chiamare  $C_i$ , conterrà tutti (e soli) i numeri interi positivi con cifra delle unità  $i$ . Ci vuol poco a convincersi che anche questa suddivisione degli interi positivi è compatibile con l’addizione e la moltiplicazione: scelti comunque due cassetti  $C_i$  e  $C_j$  e due numeri,  $a$  in  $C_i$  e  $b$  in  $C_j$ , quale sia il cassetto che contiene la somma  $a + b$  e quale sia il cassetto che contiene il prodotto  $ab$  dipende solo da  $i$  e da  $j$  e non cambia se sostituiamo  $a$  con un qualsiasi altro numero  $a'$  in  $C_i$  e  $b$  con un qualsiasi altro  $b'$  in  $C_j$ . Ad esempio, comunque scegliamo un  $a$  nel cassetto  $C_3$  ed un  $b$  nel cassetto  $C_4$ ,  $a + b$  sarà nel cassetto  $C_7$  e  $ab$  nel cassetto  $C_2$  (provare per credere).

È possibile generalizzare questa idea; vediamo come. Il cassetto  $C_0$  è costituito, abbiamo detto, dai numeri interi positivi con cifra delle unità zero. Ma questi sono precisamente i multipli di 10. I numeri nel cassetto  $C_1$ , quelli con ultima cifra 1, sono precisamente (tra gli interi positivi) quelli che nella divisione per 10 hanno resto 1. In generale, possiamo facilmente constatare che, se  $i$  è una qualsiasi delle cifre  $0, 1, 2, \dots, 9$  (che sono, guarda caso, i possibili resti nella divisione di un intero per 10) il cassetto  $C_i$

UNIVERSITÀ DEGLI STUDI DI NAPOLI “FEDERICO II”, DIPARTIMENTO DI MATEMATICA E APPLICAZIONI “R. CACCIOPPOLI”,  
VIA CINTIA — MONTE S. ANGELO, I-80126 NAPOLI, ITALY,  
e-mail: [cutolo@unina.it](mailto:cutolo@unina.it)

<http://www.dma.unina.it/~cutolo/>, <http://www.dma.unina.it/~cutolo/didattica/>

conterrà tutti e soli gli interi positivi che, divisi per 10, danno resto  $i$ . Ci accorgiamo poi che anche la suddivisione degli interi tra pari e dispari risponde alla stessa logica: i numeri pari sono quelli che, divisi per 2, danno resto zero, quelli dispari sono quelli che, divisi per due danno resto uno. Una differenza è che per la suddivisione tra pari e dispari abbiamo preso in esame tutti gli interi, mentre la suddivisione “per cifra delle unità” l’abbiamo ristretta ai soli insiemi positivi. Questa differenza, vedremo, è inessenziale. L’unica altra differenza è che la prima suddivisione può essere descritta con riferimento alla divisione per 2, nel secondo alla divisione per 10. Ma se qualcosa funziona allo stesso modo per 2 e per 10, magari si può pensare che funzioni allo stesso modo anche per altri numeri. Difatti è così. Fissiamo un arbitrario intero positivo  $m$  e disponiamo tutti gli interi (non solo quelli positivi) in una cassetiera con  $m$  cassetti, etichettati con i numeri  $0, 1, 2, \dots, m - 1$ , mettendo nel cassetto con etichetta  $i$  tutti e soli gli interi che, divisi per  $m$ , danno resto  $i$ . In questo modo otteniamo che ciascun numero intero è finito in un cassetto (ed uno solo), perché i numeri con cui abbiamo etichettato i cassetti sono precisamente i possibili resti nella divisione di un intero per  $m$ .

Queste cassettiere con  $m$  cassetti sono dunque una versione più generale dei due esempi precedenti, ed hanno anche esse, per qualsiasi valore dell’intero positivo  $m$ , la stessa proprietà di compatibilità rispetto sia all’addizione che alla moltiplicazione tra numeri interi. Questo lo verificheremo nella prossima sezione, ma discutiamo da subito di cosa ciò significhi: nella sostanza, che possiamo definire un’operazione di addizione ed una di moltiplicazione tra i “cassetti”. Consideriamo infatti fissato l’intero positivo  $m$ . Se  $A$  e  $B$  sono due dei cassetti (uguali o diversi tra loro, non importa) della nostra cassetiera con  $m$  cassetti,  $a$  è un numero in  $A$  e  $b$  un numero in  $B$ , chiamiamo  $A + B$  il cassetto a cui appartiene  $a + b$ . La definizione ha senso proprio per via della compatibilità:  $A + B$  non dipende dalla scelta di  $a$  in  $A$  e di  $b$  in  $B$ ; anche se sostituiamo  $a$  con un altro elemento  $a'$  di  $A$  e  $b$  con un altro elemento  $b'$  di  $B$  non cambia il “cassetto somma”  $A + B$ , perché  $a' + b'$  è nello stesso cassetto di  $a + b$ . Simile discorso vale per la moltiplicazione: la compatibilità garantisce che sia univocamente definito il “cassetto prodotto”  $AB$  come quel cassetto a cui appartiene  $ab$ . Abbiamo così un “ambiente di calcolo” (i matematici parlano di *struttura algebrica*): un insieme in cui siano definite delle operazioni, che benché sia definito a partire dai numeri interi non è più quello dei numeri interi: continuando nella metafora, questo ambiente è la nostra cassetiera. Anzi, abbiamo infiniti ambienti di calcolo, uno per ogni scelta dell’intero positivo  $m$ , tutti diversi tra loro e diversi dall’ambiente originale (quello dei numeri interi). L’aritmetica modulare si può descrivere come l’algebra di questi nuovi ambienti di calcolo o, per dirla in modo più educato, di queste strutture e delle loro proprietà.

## 2. IN MODO UN PO’ PIÙ PRECISO . . .

Indichiamo, come si fa di consueto, con  $\mathbb{Z}$  l’insieme dei numeri interi. Come sappiamo, se  $u, v \in \mathbb{Z}$  (cioè: se  $u$  e  $v$  sono numeri interi), dire che  $u$  divide  $v$  (o, equivalentemente, che  $v$  è divisibile per  $u$ , o ancora che  $v$  è multiplo di  $u$ ) significa dire che esiste  $k \in \mathbb{Z}$  tale che  $v = uk$ . Aggiungiamo una nuova definizione: se  $m$  è un intero positivo e  $a, b \in \mathbb{Z}$ , diciamo che  $a$  e  $b$  sono *congrui modulo  $m$* , e scriviamo in questo caso  $a \equiv_m b$ , se e solo se  $m$  divide la differenza  $a - b$ . Ad esempio,  $8 \equiv_3 2$  (perché 3 divide  $8 - 2 = 6$ ) e  $7 \equiv_{10} -3$  (perché 10 divide  $7 - (-3) = 10$ ). Vediamo alcune proprietà essenziali di questa relazione di congruenza modulo  $m$ , senza entrare troppo in dettagli (né in dimostrazioni), che possono comunque essere trovati facilmente altrove da chi lo desidera.

Innanzitutto, la congruenza modulo  $m$  è una *relazione di equivalenza* in  $\mathbb{Z}$ , nel senso che verifica le proprietà riflessiva (si ha  $a \equiv_m a$  per ogni  $a \in \mathbb{Z}$ , dal momento che  $m$  certamente divide  $0 = a - a$ ), simmetrica (se  $a$  e  $b$  sono interi tali che  $a \equiv_m b$ , allora  $b \equiv_m a$ : infatti se  $a - b$  è un multiplo di  $m$  allora anche  $b - a$ , che è l’opposto di  $a - b$ , è un multiplo di  $m$ ) e transitiva (se  $a, b, c \in \mathbb{Z}$  e si ha  $a \equiv_m b$  e  $b \equiv_m c$ , allora  $a \equiv_m c$ , infatti se  $a - b$  e  $b - c$  sono entrambi multipli di  $m$ , allora anche la loro somma  $(a - b) + (b - c) = a - c$  è un multiplo di  $m$ )<sup>1</sup>. Per ogni  $a \in \mathbb{Z}$ , l’insieme di tutti i numeri interi che siano congrui ad  $a$  modulo  $m$  si chiama *classe di resto* di  $a$  modulo  $m$  e si indica con  $[a]_m$ .<sup>2</sup> Da quali interi è costituito questo insieme? Si ha:

$$[a]_m = \{a + mk \mid k \in \mathbb{Z}\}$$

---

<sup>1</sup>useremo spesso questa osservazione, che è tanto importante quanto ovvia: se  $u$  e  $v$  sono due multipli dello stesso intero  $m$ , allora anche la loro somma è un multiplo di  $m$ ; infatti se  $u = mh$  e  $v = mk$ , per opportuni interi  $h$  e  $k$ , allora  $u + v = m(h + k)$ .

<sup>2</sup>chi ha familiarità con il linguaggio e la teoria delle relazioni di equivalenza si accorgerà che  $[a]_m$  non è altro che la classe di equivalenza di  $a$  rispetto alla relazione (di equivalenza) di congruenza modulo  $m$ .

vale a dire: gli interi congrui ad  $a$  modulo  $m$  sono quelli della forma  $a + mk$  per un opportuno  $k \in \mathbb{Z}$ . Infatti, se  $b \equiv_m a$  allora  $b - a$  è un multiplo di  $m$ , quindi  $b - a = mk$  per un opportuno  $k \in \mathbb{Z}$ , vale a dire:  $b = a + mk$ . Viceversa, se  $b$  è un numero della forma  $b = a + mk$ , per un  $k \in \mathbb{Z}$ , allora  $b - a = mk$  è multiplo di  $m$  e quindi  $b \equiv_m a$ . Questo prova l'uguaglianza  $[a]_m = \{a + mk \mid k \in \mathbb{Z}\}$ . Ad esempio, la classe di resto di 2 modulo 6, cioè  $[2]_6$  è costituita da tutti gli interi della forma  $2 + 6k$  al variare di  $k \in \mathbb{Z}$ ; per  $k$  uguale a 0, 1, 2, 3, etc. otteniamo così  $2 = 2 + 6 \cdot 0$ ,  $8 = 2 + 6 \cdot 1$ ,  $14 = 2 + 6 \cdot 2$ ,  $20 = 2 + 6 \cdot 3$ , e così via, scegliendo per  $k$  valori negativi otteniamo invece  $-4 = 2 + 6 \cdot (-1)$ ,  $-10 = 2 + 6 \cdot (-2)$ ,  $-16 = 2 + 6 \cdot (-3)$ , e così via. Abbiamo così:

$$[2]_6 = \{\dots, -22, -16, -10, -4, 2, 8, 14, 20, 26, \dots\};$$

possiamo visualizzare la classe di resto di 2 modulo 6 come l'insieme costituito dai numeri che incontriamo partendo da 2 ed percorrendo l'intera lista dei numeri interi facendo passi di lunghezza 6, sia nel verso positivo che in quello negativo. Scegliamo un qualsiasi altro numero in  $[2]_6$ , ad esempio 14. Qual è la classe di resto di 14 modulo 6? Se ci pensiamo un attimo, ci accorgiamo che è la stessa classe trovata per 2. Infatti ci possiamo spostare da 14 a 2 (con due passi “all'indietro” di lunghezza 6) e da qui, sempre con passi di lunghezza 6, raggiungere tutti i numeri in  $[2]_6$  (per esempio, siccome con cinque passi all'indietro, da 2 si raggiunge  $-28 = 2 + 6(-5)$ , si potrà raggiungere  $-28$  da 14 effettuando sette passi all'indietro; i primi due passi portano a 2, con i rimanenti cinque si arriva a  $-28$ ). Vediamo così che ogni numero in  $[2]_6$  è anche in  $[14]_6$ . Viceversa, ragionando in modo analogo ma invertendo i ruoli di 14 e 2, scopriamo che, partendo da 2 ed muovendoci solo con passi di lunghezza 6, siccome possiamo raggiungere 14 possiamo anche raggiungere ogni numero in  $[14]_6$ , quindi ogni numero in  $[14]_6$  è anche in  $[2]_6$ . Mettendo insieme questa informazione con la precedente, concludiamo che  $[14]_6 = [2]_6$ . Questa è una regola generale, si ha infatti, per ogni coppia di interi  $a$  e  $b$  e per ogni intero positivo  $m$ :

$$a \equiv_m b \iff [a]_m = [b]_m.$$

Questa è una proprietà generale delle relazioni di equivalenza, e può darsi che chi legge la abbia già incontrata. La si può verificare utilizzando la proprietà transitiva (e la proprietà simmetrica). Supponiamo infatti  $a \equiv_m b$ . Se  $c$  è un elemento di  $[b]_m$ , allora  $b \equiv_m c$  (per definizione di  $[b]_m$ ), dunque  $a \equiv_m c$  per la proprietà transitiva, vale a dire:  $c \in [a]_m$ . Questo prova l'inclusione  $[b]_m \subseteq [a]_m$ ; in modo analogo si prova l'inclusione opposta:  $[a]_m \subseteq [b]_m$ . Dunque è vero che  $[a]_m = [b]_m$  se  $a \equiv_m b$ . A guardar bene, questa dimostrazione non è altro che una versione un po' più astratta del ragionamento che abbiamo svolto sopra per arrivare all'uguaglianza  $[14]_6 = [2]_6$ . Ci resta ancora da fare una cosa: verificare l'implicazione opposta, cioè che  $a \equiv_m b$  se  $[a]_m = [b]_m$ . Ma questo è chiaro: si ha  $a \in [a]_m$ , perché  $a \equiv_m a$  (proprietà riflessiva!); dunque, se  $[a]_m = [b]_m$  allora  $a \in [b]_m$  e quindi  $a \equiv_m b$ .

Un'altra utile espressione della stessa proprietà è:

*scelta comunque una classe di resto  $A$  ed un  $a \in A$ , allora  $A = [a]_m$ .*

Le classi di resto si chiamano così perché sono strettamente legate alla nozione di resto che appare nella divisione tra numeri interi. Ricordiamo di cosa stiamo parlando:

**Teorema.** Scelti comunque due numeri interi  $a$  ed  $m$ , se  $m > 0$  esiste una ed una sola coppia ordinata  $(q, r)$  di numeri interi tali che  $a = mq + r$  e  $0 \leq r < m$ .

Come sappiamo, questi numeri  $q$  ed  $r$  si chiamano, nell'ordine, quoziante e resto nella divisione di  $a$  per  $m$ . Siccome  $a - r = mq$  è un multiplo di  $m$ , osserviamo che  $r$  è congruo ad  $a$  modulo  $m$ . Meglio ancora,  $r$  è l'unico numero congruo ad  $a$  modulo  $m$  che sia compreso tra 0 e  $m - 1$ , infatti tra gli elementi della classe  $[a]_m$  (che coincide, ricordiamo, con  $[r]_m = \{r + mk \mid k \in \mathbb{Z}\}$ , dal momento che  $r \equiv_m a$ ) quello che immediatamente segue  $r$  è  $r + m$ , che è più grande di  $m - 1$ , quello che immediatamente precede  $r$  è  $r - m$ , che è negativo. Se  $b$  è un arbitrario numero congruo ad  $a$  modulo  $m$  e  $s$  è il resto di  $b$  nella divisione per  $m$ , allora  $s \equiv_m b \equiv_m a$  e  $0 \leq s < m$ , quindi anche  $s$ , come  $r$ , è congruo ad  $a$  modulo  $m$  (proprietà transitiva!) ed è compreso tra 0 e  $m - 1$ . Ma  $r$  è l'unico numero con queste proprietà, quindi  $s = r$ . Cosa concludiamo? Che *tutti* i numeri in  $[a]_m$ , se divisi per  $m$ , hanno come resto  $r$ . Vale anche il viceversa, infatti se un numero intero  $b$ , diviso per  $m$ , ha resto  $r$ , allora  $b \equiv_m r$ ; ma  $r \equiv_m a$  e quindi  $b \equiv_m a$ , ovvero  $b \in [a]_m$ . Possiamo a questo punto dire che  $[a]_m$  è costituita da tutti e soli gli interi che, divisi per  $m$ , danno resto  $r$ . Questo spiega il nome ‘classe di resto’ dato a questi insiemi. Possiamo anche dire:

*due numeri sono congrui modulo  $m$  se e solo se hanno lo stesso resto nella divisione per  $m$ .*

Beh, allora queste classi di resto sono niente di più e niente di meno che i ‘cassetti’ che avevamo informalmente introdotto nella sezione precedente: fissato l’intero positivo  $m$ , esistono esattamente  $m$  classi di resto (di numeri interi) modulo  $m$ , una per ogni possibile resto nella divisione per  $m$ . L’insieme di queste classi di resto si indica con  $\mathbb{Z}_m$  (e si chiama l’insieme degli *interi modulo  $m$* ):

$$\mathbb{Z}_m = \{[0]_m, [1]_m, [2]_m, \dots, [m-1]_m\}$$

dove, per ciascun  $i$ ,  $[i]_m$  è l’insieme degli interi che, divisi per  $m$ , dànno resto  $i$ . Come abbiamo visto,  $[i]_m = \{i + mk \mid k \in \mathbb{Z}\}$ . Notiamo, in particolare, che  $[0]_m$  è l’insieme degli interi multipli di  $m$  (che si indica talvolta con  $m\mathbb{Z}$ ). Notiamo anche che ciascuna delle classi di resto si può rappresentare in tanti (addirittura infiniti) modi diversi, ad esempio  $[2]_6 = [14]_6 = [6000002]_6$ , quindi anche  $\mathbb{Z}_m$  si potrà descrivere in molti modi. Ad esempio,

$$\mathbb{Z}_m = \{[1]_m, [2]_m, \dots, [m-1]_m, [m]_m\},$$

dal momento che  $[0]_m = [m]_m$ , maabbiamo anche

$$\mathbb{Z}_m = \{[a]_m \mid a \in \mathbb{Z}\}.$$

Giusto per chiarire un punto lasciato in sospeso nella sezione precedente, esaminiamo il caso  $m = 10$ .  $\mathbb{Z}_{10}$  consiste delle dieci classi  $[i]_{10} = \{i + 10k \mid k \in \mathbb{Z}\}$  al variare dell’intero  $i$  tra 0 e 9. Come sono fatte queste classi di resto? Lo capiamo da un esempio: i numeri positivi in  $[7]_{10}$  sono tutti e soli quelli che hanno 7 come cifra delle unità. E quelli negativi? Si capisce che  $-7$  non è in questa classe, perché la differenza  $7 - (-7)$  è 14, che non è multiplo di 10. Invece in questa classe troviamo  $-3 = 7 - 10$ . Difatti si riconosce facilmente che  $[7]_{10}$  è costituita dagli interi positivi con ultima cifra 7 e dagli interi negativi con ultima cifra 3. È per evitare questa complicazione che, nella sezione precedente, abbiamo descritto la cassettiera a dieci cassetti limitandoci ai numeri interi positivi.

**2.1. Compatibilità.** Abbiamo detto, nella sezione precedente, che è la proprietà di compatibilità a rendere interessanti le ‘cassettiere’, perché rende possibile definire operazioni ‘tra cassetti’ in aggiunta a quelle tra numeri. Vediamo in che modo questa proprietà si può riformulare e giustificare.

**Teorema** (Compatibilità delle congruenze). *Siano dati un intero positivo  $m$  e  $a, a', b, b' \in \mathbb{Z}$ . Allora:*

$$\left. \begin{array}{l} a \equiv_m a' \\ e \\ b \equiv_m b' \end{array} \right\} \implies \left\{ \begin{array}{l} a+b \equiv_m a'+b' \\ e \\ ab \equiv_m a'b' \end{array} \right.$$

*Dimostrazione.* Supponiamo che valgano  $a \equiv_m a'$  e  $b \equiv_m b'$ . Questa ipotesi significa che  $a-a'$  e  $b-b'$  sono entrambi multipli di  $m$ . Per provare  $a+b \equiv_m a'+b'$  ci serve verificare che la differenza  $(a+b) - (a'+b')$  è un multiplo di  $m$ . Ma questo non è difficile:  $(a+b) - (a'+b') = (a-a') + (b-b')$  è la somma tra due multipli di  $m$ , quindi è essa stessa un multiplo di  $m$ . Dunque è vero che  $a+b \equiv_m a'+b'$ .

Proviamo ora che  $ab \equiv_m a'b'$ . Aggiungendo e sottraendo  $ab'$  a  $ab - a'b'$  si ha

$$ab - a'b' = ab - ab' + ab' - a'b' = a(b-b') + (a-a')b'.$$

Ora, sia  $a(b-b')$  che  $(a-a')b'$  sono multipli di  $m$ , quindi  $ab - a'b'$  è multiplo di  $m$ , sicché  $ab \equiv_m a'b'$ . A questo punto l’asserto è provato.  $\square$

Come già detto, questo risultato è il punto essenziale dell’aritmetica modulare. Lo possiamo vedere da due punti di vista.

In primo luogo, supponiamo di voler conoscere non il risultato di alcune operazioni tra interi, ma solo il resto che questo risultato ha nella divisione per un certo intero positivo  $m$ . Ad esempio, per assegnati numeri interi  $a, b, c, d, e, f$  abbiamo bisogno di calcolare il resto modulo  $m$  di  $k := (a^2b - c)(d + e^5) + f$ . Possiamo sostituire  $a, b, c, d, e, f$  con numeri interi  $a', b', c', d', e', f'$ , ciascuno congruo modulo  $m$  a quello che sostituisce ( $a \equiv_m a'$ ,  $b \equiv_m b'$  e così via) e calcolare  $k' = ((a')^2b' - c')(d' + (e')^5) + f'$ , avendo scelto naturalmente i numeri che sostituiamo in modo che questo secondo calcolo sia più semplice dell’originale. Siccome  $k' \equiv_m k$ , per la proprietà di compatibilità,  $k'$  ha lo stesso resto di  $k$  nella divisione per  $m$ . Notiamo in particolare come questo vale anche per le potenze: ad esempio,  $a^2 = a \cdot a \equiv_m a' \cdot a' = (a')^2$ . Ci si trova in questa situazione molto più frequentemente di quanto possa sembrare a prima vista, sia in contesti puramente matematici che in situazioni di applicazione della matematica a problemi del mondo reale, ad esempio, quando dobbiamo far conti che abbiano a che fare con lo scorrere del tempo, che calcoliamo ciclicamente (l’orario si azzera ogni ventiquattr’ore, il giorno delle settimana ogni sette giorni,

etc.). Vediamo subito qualche esempio di questo secondo tipo (dove l'espressione ‘mondo reale’ è usata in un senso piuttosto generoso).

**Esempio 1.** È domenica, e mio cugino Asdrubale vuole sapere da me che giorno (della settimana) sarà tra  $8500 \cdot 33 + 4^{200}$  giorni. Asdrubale non è particolarmente sensibile a quell'utile qualità chiamata buonsenso, e non accetta come valida la mia risposta che l'espressione “tra  $4^{200}$  giorni” non ha alcun significato reale: ben prima che possa essere trascorso una tale lasso di tempo non esisterà più nessuno, o almeno nessuno che misuri il tempo in giorni e in settimane come facciamo noi. Lui vuole la risposta comunque! E allora diamogliela. Si tratta innanzitutto di calcolare il resto modulo 7 di  $8500 \cdot 33 + 4^{200}$ . Procediamo con ordine:  $8500 = 8000 + 500 \equiv_7 1000 + 500$ , dal momento che  $8000 - 1000$  è chiaramente multiplo di 7. Dunque,  $8500 \equiv_7 1500$ . Ma  $1500 = 1400 + 100 \equiv_7 100$  e  $100 = 70 + 30 \equiv_7 30 = 28 + 2 \equiv_7 2$ . Dunque,  $8500 \equiv_7 2$ . Inoltre  $33 \equiv_7 -2$ . Quindi  $8500 \cdot 33 \equiv_7 2 \cdot (-2) = -4 \equiv_7 3$ . Esaminiamo ora  $4^{200}$ . Abbiamo  $4^2 = 16 \equiv_7 2$ , quindi  $4^3 = 4^2 \cdot 4 \equiv_7 2 \cdot 4 \equiv_7 8 \equiv_7 1$ . Ma allora abbiamo anche  $4^4 \equiv_7 4^3 \cdot 4 \equiv_7 4$ ,  $4^5 \equiv_7 4^3 \cdot 4^2 \equiv_7 4^2$ ,  $4^6 \equiv_7 4^3 \cdot 4^3 \equiv_7 1$ : le “potenze di 4 modulo 7” si ripetono di tre in tre, il che rende facile calcolarle. Precisiamo questa idea. Da  $4^3 \equiv_7 1$  segue  $4^{3k} = (4^3)^k \equiv_7 1^k = 1$  per ogni intero positivo  $k$ . Immaginiamo di aver diviso 200 per 3, ottenendo un quoziente  $k$  ed un resto  $r$ , dunque  $200 = 3k + r$ . Allora  $4^{200} = 4^{3k+r} = 4^{3k} \cdot 4^r \equiv_7 1 \cdot 4^r$ . Quindi per calcolare  $4^{200}$  modulo 7 l'unica cosa che serve conoscere è il resto di 200 nella divisione per 3. Esiste un metodo rapidissimo per calcolarlo, lo vedremo più avanti, ma procediamo in modo ingenuo:  $200 = 2 \cdot 100$  e  $100 \equiv_3 1$  (perché 99 è multiplo di 3); allora  $200 \equiv_3 2 \cdot 1 = 2$  e da ciò segue che 2 è il resto  $r$  cercato. Allora, per le osservazioni fatte sopra,  $4^{200} \equiv_7 4^2 \equiv_7 2$  (notare che non abbiamo avuto bisogno di calcolare il quoziente  $k$ ; che  $4^2 \equiv_7 2$  lo avevamo già osservato). In definitiva,

$$8500 \cdot 33 + 4^{200} \equiv_7 3 + 2 = 5.$$

Sappiamo allora che questo gran numero (di giorni) è un multiplo di 7 più 5. Aggiungere alla data un numero di giorni multiplo di 7 (cioè un numero intero di settimane) non cambia il giorno della settimana, quindi per rispondere alla domanda di Asdrubale basta aggiungere cinque giorni al giorno attuale (che, abbiamo detto, è domenica), ottenendo un venerdì. Possiamo dire allora ad Asdrubale che il giorno che tanto gli sta a cuore sarà un venerdì. Contento lui ...  $\square$

**Esempio 2.** La riproduzione del batterio *Duplicator Freneticus Suicidalis*, o *DFS*, (fortunatamente non particolarmente dannoso se non per se stesso) ha questo curioso andamento. In ambiente favorevole, ogni esemplare si duplica dopo esattamente tre minuti di vita, dando luogo a due individui identici, che a loro volta si duplicheranno (istantaneamente) dopo esattamente tre minuti, e così via. C'è però una restrizione: ogni ambiente chiuso (ad esempio una provetta) ha un suo limite di popolazione, diciamo di  $n$  batteri. Se dopo una duplicazione della popolazione della provetta il numero di batteri raggiunge o supera  $n$ , istantaneamente  $n$  batteri muoiono abbassando la popolazione al di sotto di  $n$  (perché? la natura è misteriosa e a volte un poco stupidita); passati tre minuti la popolazione superstite si duplicherà, come al solito. Ad esempio, se per una certa provetta questo limite è di 90 batteri, e in un certo istante la provetta contiene precisamente 50 batteri vivi, appena “nati”, dopo tre minuti questi 50 batteri si duplicheranno, ma allora diventerebbero 100, il che non è consentito dalla natura del batterio e della provetta, quindi 90 batteri muoiono e ne restano in vita 10, che dopo tre minuti diventano 20, dopo sei minuti 40, poi 80, e dopo altri tre minuti diventano 70 (dovrebbero essere 160, ma 90 di essi devono morire), e così si ripetono questi (molto) strani cicli biologici.

In un certo laboratorio è stata preparata una provetta per la quale il limite di popolazione di batteri *DFS* che non può essere raggiunto è di 85 batteri. Nell'istante  $t_0$  in cui si forma, un (solo) batterio viene lasciato cadere nella provetta (che prima non ne conteneva), in modo che dopo tre minuti si possa duplicare e dare avvio al popolamento della provetta. La questione è: quanti batteri saranno nella provetta esattamente cinque ore e dieci secondi dopo  $t_0$ ? In ogni ora avvengono  $20 (= 60/3)$  cicli di duplicazione dei batteri, quindi allo scoccare della quinta ora ne saranno avvenuti cento. Ogni duplicazione è un “raddoppio modulo 85”, vale a dire: se subito prima della duplicazione la provetta  $X$  conteneva  $n$  batteri, subito dopo ne conterrà  $2n$  (se  $2n < 85$ ) o  $2n - 85$  (se  $2n > 85$ ; ovviamente in nessun caso possiamo avere  $2n = 85$ , che causerebbe l'estinzione della popolazione dei batteri in  $X$ ); in ogni caso il numero dei batteri è il resto nella divisione di  $2n$  per 85. Dopo cento duplicazioni dei batteri il numero sarà allora il resto di  $2^{100}$  nella divisione per 85 (dobbiamo moltiplicare cento volte per due il batterio iniziale, ma la moltiplicazione è fatta “modulo 85”). Calcoliamolo. La prima potenza di 2 che si avvicini al modulo 85 è  $2^6 = 64 = 85 - 21 \equiv_{85} -21$ . Allora  $2^7 \equiv_{85} 2(-21) = -42$  e  $2^8 \equiv_{85} 2(-42) = -84 \equiv_{85} 1$ . Ragionando come

fatto nell'esempio precedente, allora ci basta calcolare il resto modulo 8 dell'esponente 100 per calcolare  $2^{100}$  modulo 85. Abbiamo  $100 \equiv_8 20 \equiv_8 4$ , quindi questo resto è 4. Il numero dei batteri cercato sarà allora congruo, modulo 85, a  $2^4 = 16$  (posto  $100 = 8k+4$ , abbiamo infatti  $2^{100} = (2^8)^k \cdot 2^4 \equiv_{85} 1^k \cdot 2^4 = 2^4$ ), dunque proprio 16.  $\square$

**Esempio 3.** Ci troviamo nella stessa situazione dell'esempio precedente, ma stiamo conducendo un esperimento parallelo in una seconda provetta  $Y$ , più piccola di  $X$ , per la quale il limite di popolazione di batteri  $DFS$  che non può essere raggiunto è di 30 batteri. Come nel caso di  $X$ , al tempo  $t_0$  la provetta  $Y$  contiene esattamente un batterio, neonato, e vogliamo conoscere il numero di batteri in  $Y$  cinque ore e pochi secondi dopo  $t_0$ , vale a dire: dopo cento duplicazioni. Dobbiamo allora calcolare il resto di  $2^{100}$  nella divisione per 30. Possiamo provare a ragionare come fatto nei due esempi precedenti. Una potenza di 2 che si avvicina molto a 30 è  $2^5 = 32$ . Abbiamo  $2^5 \equiv_{30} 2$ , quindi  $2^6 \equiv_{30} 2^2$ ,  $2^7 \equiv_{30} 2^3$ ,  $2^8 \equiv_{30} 2^4$ ,  $2^9 \equiv_{30} 2^5 \equiv_{30} 2$ ,  $2^{10} \equiv_{30} 2^2$  e così via: le potenze di 2 modulo 30 si ripetono di quattro in quattro a partire da  $2^1 = 2$ . Precisiamo questa frase: supponiamo che  $s$  sia un intero *positivo*. Allora:  $2^s = 2 \cdot 2^{s-1} \equiv_{30} 2^5 \cdot 2^{s-1} = 2^{s+4}$ . Ora, per ogni intero positivo  $n$ , se  $q$  ed  $r$  sono rispettivamente il quoziente ed il resto nella divisione di  $n-1$  per 4, abbiamo  $n-1 = 4q+r$ , quindi  $n = 4q+(r+1)$ . Dunque  $n$  si ottiene aggiungendo un certo numero ( $q$ , ma non importa) di volte 4 all'intero *positivo*  $r+1$ . Per quanto visto sopra, questo garantisce che  $2^{r+1} \equiv_{30} 2^{r+1+4} \equiv_{30} 2^{r+1+4 \cdot 2} \equiv_{30} 2^{r+1+4 \cdot 3} \equiv_{30} \dots \equiv_{30} 2^{r+1+4q} = 2^n$ .

Ora sappiamo come calcolare la nostra potenza  $2^{100}$  modulo 30. Il resto di  $99 = 100 - 1$  modulo 4 è 3, perché 100 è multiplo di 4, quindi  $99 \equiv_4 -1 \equiv_4 3$  (con riferimento alle notazioni appena usate, 100 è  $n$ , dunque  $r$  è 3). Allora  $2^{100} \equiv 2^{3+1} = 16$ ; la provetta, al tempo indicato (cinque ore dopo  $t_0$ ) conterrà 16 batteri.

Osserviamo come, in questo caso, il calcolo delle potenze è un po' più complicato che nel caso dei due esempi precedenti. La differenza è che mentre le potenze di 4 modulo 7 e quelle di 2 modulo 85 si ripetono a partire da quella di esponente zero ( $1 = 4^0 \equiv_7 4^3$  e  $1 = 2^0 \equiv_{85} 2^8$ ), quelle di 2 modulo 30 si ripetono anch'esse, ma non da quella di esponente 0. Infatti, per ogni intero positivo  $t$  si ha  $1 = 2^0 \not\equiv_{30} 2^t$ , perché per ogni numero congruo a 1 modulo 30 è dispari (dal momento che è  $30k+1$  per qualche  $k \in \mathbb{Z}$ ), mentre tutte le potenze di 2 con esponente intero positivo sono, ovviamente, pari. Come abbiamo visto, si ha una ripetizione periodica delle potenze di 2 modulo 30 solo a partire dalla potenza di esponente uno, cioè da  $2 = 2^1 = 2^5 = 2^9 \dots$ .  $\square$

Fissiamo in forma generale una osservazione sul calcolo di potenze in aritmetica modulare. Siano  $m$  (come al solito) un intero positivo e  $a$  un intero arbitrario. Come abbiamo visto negli esempi, può capitare che esista un intero positivo  $t$  tale che  $a^t \equiv_m 1$ . Limitandoci al caso più semplice, supponiamo che questo accada (a titolo di notizia: si verifica questo caso se e solo se  $a$  ed  $m$  sono coprimi<sup>3</sup>) In questo caso il minimo tale  $t$  si chiama *periodo* (o anche periodo moltiplicativo) di  $a$  modulo  $m$ , o anche, in modo più appropriato, periodo (moltiplicativo) di  $[a]_m$ . Estendendo al caso generale i ragionamenti svolti negli esempi, si ha che per ogni intero  $n \geq 0$  vale  $a^n \equiv_m a^r$ , dove  $r$  è il resto nella divisione di  $n$  per  $t$ .

**2.2. Le operazioni in  $\mathbb{Z}_m$ .** Veniamo al secondo punto di vista. Come visto già parlando di "cassettiere" la proprietà di compatibilità rende possibile definire operazioni di addizione e moltiplicazione tra "cassetti". La versione "precisata" di questa idea è che la proprietà di compatibilità garantisce la possibilità di definire, per ogni fissato intero  $m$ , operazioni di addizione e moltiplicazione tra classi di resto modulo  $m$ , cioè tra elementi di  $\mathbb{Z}_m$ . Si usano per queste operazioni gli stessi simboli + e  $\cdot$  che usiamo per le corrispondenti operazioni tra numeri; le nuove operazioni sono definite da:

$$[a]_m + [b]_m = [a + b]_m \quad \text{e} \quad [a]_m \cdot [b]_m = [ab]_m$$

per ogni  $a, b \in \mathbb{Z}$ . Cosa significa? Niente di diverso di quanto abbiamo detto parlando di cassetti nella sezione introduttiva: la somma tra due classi di resto  $A$  e  $B$  (modulo  $m$ ) si ottiene prendendo un numero  $a$  in  $A$  ed uno  $b$  in  $B$ , sommando questi due numeri e considerando come risultato la classe a cui appartiene  $a + b$ . Infatti, come abbiamo visto sopra, se  $a \in A$  e  $b \in B$  allora  $A = [a]_m$  e  $B = [b]_m$ , e la classe a cui appartiene  $a + b$  è  $[a + b]_m$ . Ripetiamolo ancora una volta: questa "classe somma" non dipende dalla scelta di quale particolare elemento  $a$  abbiamo selezionato in  $A$  e quale  $b$  in  $B$ : se  $a' \in [a]_m$  e  $b' \in [b]_m$  allora  $a \equiv_m a'$  e  $b \equiv_m b'$ , quindi, per il teorema sulla compatibilità,  $a + b \equiv_m a' + b'$ , cioè  $[a' + b']_m = [a + b]_m$ . Analogi discorsi vale per la moltiplicazione.

<sup>3</sup>due numeri interi sono coprimi, o *primi tra loro*, o *relativamente primi*, se e solo se hanno 1 come massimo comune divisore. In modo equivalente, due interi sono coprimi se e solo se non esiste alcun numero primo che li divida entrambi.

Abbiamo ora, per ogni fissato  $m$ , un insieme,  $\mathbb{Z}_m$ , e due operazioni definite tra elementi di  $m$ ; abbiamo dunque quella che in matematica si chiama una struttura algebrica. Queste operazioni verificano le consuete proprietà (commutativa, associativa, distributiva) che ci sono familiari dall'aritmetica elementare, cioè le stesse proprietà delle operazioni tra numeri interi. Senza entrare in dettagli diciamo che queste strutture algebriche sono dello stesso tipo di quelle nelle quali siamo abituati a fare i conti, in  $\mathbb{Z}$ , nell'insieme dei numeri razionali, nell'insieme dei numeri reali. In algebra queste strutture si chiamano anelli, più precisamente, nel nostro caso, anelli commutativi unitari.

È utile visualizzare queste operazioni nelle cosiddette tavole di Cayley, di cui qui vediamo alcuni esempi, nei casi  $m = 5$  e  $m = 6$ . Iniziamo con  $\mathbb{Z}_5$ :

$+$	0	1	2	3	4	$\cdot$	0	1	2	3	4
$\mathbb{Z}_5 :$	0	0	1	2	3	4	0	0	0	0	0
	1	1	2	3	4	0	1	0	1	2	3
	2	2	3	4	0	1	2	0	2	4	1
	3	3	4	0	1	2	3	0	3	1	4
	4	4	0	1	2	3	4	0	4	3	2

La prima tabella descrive l'addizione, la seconda la moltiplicazione in  $\mathbb{Z}_5$ . In entrambe, per motivi di leggibilità, abbiamo scritto 0, 1, 2, 3, 4 piuttosto che  $[0]_5, [1]_5, [2]_5, [3]_5, [4]_5$ . Come sono composte e come vanno lette queste tabelle? Molto semplice: nella prima tabella, incrociando una riga, intestata da  $i$ , con una colonna, intestata da  $j$ , otteniamo un numero che rappresenta la classe  $[i + j]_5$ . Ad esempio, la terza riga (quella intestata da 2) ha uno zero nella quarta posizione (colonna intestata da 3). Questo esprime il fatto che  $[2]_5 + [3]_5 = [0]_5$ . La seconda tabella è costruita allo stesso modo ma con riferimento alla moltiplicazione: nella stessa posizione in cui prima abbiamo trovato uno zero abbiamo ora 1, perché  $[2]_5 \cdot [3]_5 = [1]_5$ . La prima riga (intestata da 0) della tabella per l'addizione riproduce la riga delle intestazioni, perché descrive le somme tra  $[0]_5$  e gli elementi di  $\mathbb{Z}_5$ , che coincidono con gli elementi stessi:  $[0]_5 + [j]_5 = [j]_5$  per ogni scelta di  $j$ ; questa proprietà si esprime dicendo che  $[0]_5$  è *elemento neutro* per l'operazione  $+$ . Lo stesso vale per la seconda riga (quella intestata da 1) della tabella per la moltiplicazione:  $[1]_5$  è elemento neutro per l'operazione  $\cdot$  in  $\mathbb{Z}_5$ .

Notiamo la simmetria delle tabelle rispetto alla diagonale dall'angolo in alto a sinistra a quello opposto. Sia ha questa simmetria precisamente perché le operazioni considerate verificano la proprietà commutativa; chi legge sa riconoscere il perché?

Confrontiamo le tavole di Cayley per  $\mathbb{Z}_5$  con quelle per  $\mathbb{Z}_6$ :

$+$	0	1	2	3	4	5	$\cdot$	0	1	2	3	4	5
$\mathbb{Z}_6 :$	0	0	1	2	3	4	5	0	0	0	0	0	0
	1	1	2	3	4	5	0	1	0	1	2	3	4
	2	2	3	4	5	0	1	0	2	4	0	2	4
	3	3	4	5	0	1	2	0	3	0	3	0	3
	4	4	5	0	1	2	3	0	4	2	0	4	2
	5	5	0	1	2	3	4	0	5	4	3	2	1

Vediamo che mentre la tavola per l'addizione è molto simile a quella che abbiamo costruito per  $\mathbb{Z}_5$ , quella relativa alla moltiplicazione è molto diversa. Nel caso di  $\mathbb{Z}_5$ , esclusa la riga intestata da zero, che contiene solo zeri, in ogni riga sono rappresentati (una ed una sola volta) tutti gli elementi di  $\mathbb{Z}_5$ . Nel caso di  $\mathbb{Z}_6$ , invece, questo accade solo per la seconda e l'ultima riga, quelle intestate da 1 e 5, non per le altre. Questa differenza è molto significativa: ci dice che  $\mathbb{Z}_5$  e  $\mathbb{Z}_6$  sono strutture algebriche molto diverse tra loro. In  $\mathbb{Z}_5$  ogni elemento diverso dallo zero è, come si dice, *invertibile*, cioè moltiplicato per un opportuno elemento dà  $[1]_5$ , che è, come accennato sopra, elemento neutro rispetto alla moltiplicazione in  $\mathbb{Z}_5$ . Questo si vede dalla tavola di Cayley per la moltiplicazione di  $\mathbb{Z}_5$ : in ogni riga, esclusa quella intestata da 0, appare 1. Ad esempio, nella riga intestata da 2 abbiamo trovato 1 in corrispondenza della colonna intestata da 3, perché  $[2]_5 \cdot [3]_5 = [1]_5$ ; possiamo esprimere questo fatto dicendo, appunto, che  $[2]_5$  è invertibile in  $\mathbb{Z}_5$  e che  $[3]_5$  è inverso di  $[2]_3$ . In questo  $\mathbb{Z}_5$  si comporta come l'anello  $\mathbb{Q}$  dei numeri razionali o quello  $\mathbb{R}$  dei numeri reali (anelli di questo tipo si chiamano *campi*), in cui tutti gli elementi diversi da zero hanno inverso (il reciproco) o, equivalentemente, in cui è sempre possibile dividere un elemento per un elemento diverso da zero, e non come  $\mathbb{Z}$ , in cui queste proprietà non valgono. Al contrario,  $\mathbb{Z}_6$  ha proprietà più deboli di quelle che valgono in  $\mathbb{Z}$ . Ad esempio, in  $\mathbb{Z}$  (come in  $\mathbb{Q}$  ed in  $\mathbb{R}$ , ed anche in  $\mathbb{Z}_5$ ) vale la legge di annullamento

del prodotto: se il prodotto tra due numeri è zero, uno dei due numeri deve essere zero. Questo non vale in  $\mathbb{Z}_6$ , come si vede anche dalla tabella; ad esempio  $[2]_6 \cdot [3]_6 = [0]_6$ , benché  $[2]_6 \neq [0]_6 \neq [3]_6$ .

Vediamo quindi che  $\mathbb{Z}$ ,  $\mathbb{Z}_5$  e  $\mathbb{Z}_6$  sono strutture algebriche di natura molto diversa tra loro. Senza entrare in dettagli che qui saranno fuori luogo, diciamo che la natura dei singoli anelli  $\mathbb{Z}_m$  è determinata dalle proprietà aritmetiche dei numeri  $m$  che li definiscono. Ad esempio, si dimostra che  $\mathbb{Z}_m$  è un campo se e solo se  $m$  è un numero primo, su questo torneremo [più avanti](#).

**2.3. Approfondimenti.** In queste prime due sezioni abbiamo menzionato due nozioni, quella di *partizione* e quella di *relazione di equivalenza*, che sono tra le più fondamentali dell'intera matematica. La seconda rientra nei programmi scolastici standard ed è generalmente nota, la prima forse no. Definiamola: se  $A$  è un insieme, una partizione di  $A$  è un insieme  $F$  di sottoinsiemi non vuoti di  $A$  con la proprietà che ogni elemento di  $A$  appartenga ad uno ed un solo elemento di  $F$ . Esempi ne abbiamo già visti nella sezione iniziale: se  $P$  e  $D$  sono l'insieme dei numeri interi pari e l'insieme dei numeri interi dispari, allora l'insieme  $\{P, D\}$  è una partizione di  $\mathbb{Z}$ ; se  $\mathbb{Z}^+$  e  $\mathbb{Z}^-$  sono l'insieme dei numeri interi positivi e l'insieme dei numeri interi negativi, allora anche  $\{\mathbb{Z}^-, \{0\}, \mathbb{Z}^+\}$  è una partizione di  $\mathbb{Z}$ .

Se  $\sim$  è una relazione di equivalenza in  $A$  e  $a \in A$ , si chiama classe di equivalenza di  $a$  rispetto a  $\sim$  l'insieme di tutti gli elementi di  $A$  che sono equivalenti (rispetto alla relazione  $\sim$ ) ad  $a$ . Ad esempio, come [già detto](#), le classi di resto sono classi di equivalenza. L'insieme delle classi di equivalenza degli elementi di  $A$  rispetto a  $\sim$ , che si chiama *insieme quoziante* e si indica con  $A/\sim$  è (lo si dimostra) una partizione di  $A$ . Viceversa, per ogni partizione  $F$  di  $A$  esiste una ed una sola relazione di equivalenza  $\sim$  tale che  $S/\sim$  sia proprio  $F$ . Si ha così una corrispondenza biettiva tra l'insieme di tutte le relazioni di equivalenza in  $A$  e quello di tutte le partizioni di  $A$ . Questa perfetta corrispondenza fa sì che studiare le relazioni di equivalenza in  $A$  sia essenzialmente lo stesso che studiare le partizioni in  $A$  e tutto quello che diciamo a proposito di partizioni si può riformulare in termini di relazioni di equivalenza, e viceversa. È per questo motivo che abbiamo potuto presentare la nozione di compatibilità prima riferendoci (anche se informalmente) a partizioni di  $\mathbb{Z}$ , poi a relazioni di equivalenza. A questo punto a chi legge non sfuggirà che, per ogni  $m$ , l'insieme  $\mathbb{Z}_m$  che abbiamo definito in questa sezione non è niente altro che l'insieme quoziante di  $\mathbb{Z}$  rispetto alla relazione di congruenza modulo  $m$ .

Tornando alla questione della compatibilità, si può dimostrare che le nostre relazioni di congruenza modulo  $m$  e la relazione di uguaglianza (che, come stiamo per vedere, si può fare rientrare nello stesso discorso) sono le sole relazioni di equivalenza compatibili con l'addizione in  $\mathbb{Z}$ : non ce ne sono altre. Ce ne sono invece di altre che siano compatibili con la moltiplicazione—un esempio l'abbiamo già fatto [all'inizio di questo testo](#). Chi legge si può divertire a cercare altre partizioni di  $\mathbb{Z}$  che siano compatibili con la moltiplicazione o che non lo siano. Una del primo tipo è la partizione che consiste di  $\{0\}$ ,  $\{1, -1\}$ , l'insieme dei numeri primi e dei loro opposti, l'insieme di tutti gli altri numeri interi (i numeri composti); una del secondo è la partizione in numeri che siano quadrati di interi  $\{0, 1, 4, 9, \dots\}$  e numeri che non siano.

Infine, giusto per non lasciarlo non detto, è il caso di avvertire che (quasi) tutto ciò che abbiamo detto facendo riferimento ad un intero positivo  $m$  ha senso e continua a valere per *tutti* gli interi: la restrizione ai positivi serve solo a semplificare l'esposizione. Ad esempio, si definisce la relazione di congruenza modulo un intero arbitrario  $m$  esattamente come fatto per gli interi positivi. Però non c'è nessun reale vantaggio nel farlo: si vede molto rapidamente che due interi sono congrui modulo 0 se e solo se coincidono (quindi la relazione di congruenza modulo 0 è la relazione di uguaglianza), e sono congrui modulo un intero negativo  $m$  se e solo se lo sono modulo l'intero positivo  $-m$  (quindi le relazioni di congruenza modulo i numeri negativi coincidono con quelle modulo i numeri positivi; non ricaviamo niente di nuovo se li aggiungiamo al discorso). Anche la divisione con resto, quella richiamata nel precedente [teorema](#), si può effettuare anche con divisori  $m$  negativi, ma (ed a questo si riferisce il “quasi” di qualche rigo fa) non per zero: una versione più generale del teorema si ottiene sostituendo l'ipotesi ‘ $m > 0$ ’ con ‘ $m \neq 0$ ’ e la restrizione ‘ $0 \leq r < m$ ’ con ‘ $0 \leq r < |m|$ ’.

### 3. CRITERI DI DIVISIBILITÀ

Ci occuperemo qui di uno dei grandi misteri dell'istruzione matematica nella scuola. A tutti noi sono stati insegnati, già nella scuola elementare, i cosiddetti ‘criteri di divisibilità’ per alcuni numeri: 2, 3, 5, 11. Sono dei semplicissimi metodi che permettono di stabilire rapidamente se un assegnato numero intero positivo è o meno divisibile per, appunto, 2, 3, 5, 11. Ma cosa ci assicura che questi criteri forniscano sempre risposte corrette? La risposta si trova nell'aritmetica modulare.

Questi criteri sono riferiti alla scrittura di un numero intero in base 10 (che è quella che abitualmente usiamo). Ricordiamo di cosa si tratta, partendo da un esempio. Se scriviamo  $n = 2375$  stiamo dicendo che  $n$  è la somma  $5 + 7 \cdot 10 + 3 \cdot 10^2 + 2 \cdot 10^3$  (e chiamiamo 5, 7, 3, 2, nell'ordine, cifra delle unità, delle decine, delle centinaia, delle migliaia). In generale, utilizziamo una stringa di cifre  $\langle a_t a_{t-1} a_{t-2} \dots a_2 a_1 a_0 \rangle$  (dove  $t$  è un numero intero non negativo, per cifra intendiamo un numero intero compreso tra 0 e 9, ed usiamo il simbolo  $\langle \dots \rangle$  all'unico scopo di non confondere, come sarebbe altrimenti possibile, questa scrittura con quella del prodotto tra le cifre<sup>4</sup>) per indicare il numero  $a_0 + 10a_1 + 10^2a_2 + \dots + 10^t a_t$ . In forma più compatta:

$$\langle a_t a_{t-1} a_{t-2} \dots a_2 a_1 a_0 \rangle = \sum_{i=0}^t a_i 10^i.$$

Poniamo  $n = \langle a_t a_{t-1} a_{t-2} \dots a_2 a_1 a_0 \rangle$  e cerchiamo di calcolare  $n$  modulo  $m$ , per alcuni valori dell'intero  $m$ . È abbastanza ovvio che

$$n \equiv_{10} a_0,$$

dunque  $a_0$  (la cifra delle unità) è proprio il resto di  $n$  nella divisione per 10. Fermiamoci un attimo per una semplice osservazione:

**Lemma.** Siano  $u, v \in \mathbb{Z}$  e sia  $m$  un intero positivo. Se  $u \equiv_m v$ , allora  $u \equiv_d v$  per ogni divisore positivo  $d$  di  $m$ ,

*Dimostrazione.* Se  $u \equiv_m v$ , allora  $u - v$  è un multiplo di  $m$ . Se  $d$  è un divisore di  $m$ , tutti i multipli di  $m$  sono anche multipli di  $d$ , quindi  $u - v$  è un multiplo di  $d$ , vale a dire:  $u \equiv_d v$ .  $\square$

Ne deduciamo che, con le notazioni usate sopra, in conseguenza di  $n \equiv_{10} a_0$  abbiamo anche  $n \equiv_5 a_0$  e  $n \equiv_2 a_0$ . Da ciò i ben noti criteri di divisibilità per 2, per 5 e per 10: un intero positivo  $n$  è divisibile per 2 (risp. 5, 10) se e solo se lo è la sua cifra delle unità. Detto diversamente: un intero positivo è pari se e solo se la sua cifra delle unità è pari, è divisibile per 5 se e solo se la sua cifra delle unità è una tra 5 e 0 (questo perché tra le cifre solo queste due sono numeri divisibili per 5), è divisibile per 10 se e solo se la sua cifra delle unità è 0.

Cosa succede per altre potenze di 10? Abbiamo sicuramente  $n \equiv_{100} 10a_1 + a_0 = \langle a_1 a_0 \rangle$ , dal momento che nella somma  $\sum_{i=0}^t a_i 10^i$  ogni addendo  $10^i a_i$  per  $i > 1$  è multiplo di 100. Dunque, modulo 100, il nostro intero positivo  $n$  è congruo al numero formato dalle sue ultime due cifre. Usando il lemma precedente, otteniamo anche che vale l'analogia congruenza modulo un qualsiasi divisore di 100, quindi  $n \equiv_d \langle a_1 a_0 \rangle$  se  $d$  è uno tra 4, 20, 25, 50, 100. Per ciascuno di questi numeri abbiamo dunque il criterio di divisibilità:  $n$  è divisibile per  $d$  se e solo se lo è il numero costituito dalle ultime sue due cifre. Si potrebbe continuare all'infinito considerando tutte le potenze di 10 ed i loro divisori ed ottenendo per essi criteri di divisibilità: per ogni intero positivo  $\ell$ , infatti,  $n$  è congruo modulo  $10^\ell$  al numero formato dalle sue ultime  $\ell$  cifre (questo numero è il resto di  $n$  nella divisione per  $10^\ell$ ).

Veniamo ad un caso più interessante: ragioniamo modulo 9. Abbiamo  $10 \equiv_9 1$  e quindi  $10^i \equiv_9 1^i \equiv_9 1$  per ogni intero non negativo  $i$ . Allora

$$n = \langle a_t a_{t-1} a_{t-2} \dots a_2 a_1 a_0 \rangle = \sum_{i=0}^t a_i 10^i \equiv_9 \sum_{i=0}^t a_i.$$

Abbiamo provato che ogni numero intero positivo è congruo, modulo 9, alla somma delle sue cifre. Ne ricaviamo un metodo per calcolare il resto modulo 9 di un arbitrario intero positivo: sostituire ripetutamente al numero la somma delle sue cifre sino ad ottenere un numero di una sola cifra; questo numero è il resto cercato se è diverso da 9, se invece questo numero è 9 allora il resto è 0. Vediamo un esempio: sia  $n = 672455978913$ . Allora  $n \equiv_9 6 + 7 + 2 + 4 + 5 + 5 + 9 + 7 + 8 + 9 + 1 + 3$ . Notiamo che non è necessario effettuare realmente la somma, dato che siamo interessati solo al calcolo modulo 9. Possiamo dunque cancellare i 9, 2 con uno dei 7, 6 con 3, 4 con uno dei 5 e 8 con 1, ottenendo  $n \equiv_9 5 + 7 = 12 \equiv_9 1 + 2 = 3$ . Il resto, dunque è 3. Partendo da 1008 otteniamo invece  $1008 \equiv_9 1 + 0 + 0 + 8 = 9 \equiv_9 0$ ; ovviamente il resto non è 9 ma 0.

Ricaviamo poi, per il solito lemma, che ogni intero positivo è congruo alla somma delle sue cifre anche modulo 3. Da queste considerazioni seguono (come caso particolare, il nostro risultato dice qualcosa in più) i criteri di divisibilità per 9 e per 3: un numero intero è divisibile per 9, ovvero per 3 se e solo

<sup>4</sup>per intenderci meglio, anche alla luce di quanto stiamo per dire: se  $t = 1$ ,  $a_0 = 2$  e  $a_1 = 7$ , se scriviamo  $\langle a_1 a_0 \rangle$  intendiamo il numero 72, non il numero  $14 = 7 \cdot 2$ , come invece siamo portati a fare quando scriviamo  $a_1 a_0$ .

se lo è la somma delle sue cifre ed anche un altro classico ed un po' misterioso strumento che viene spesso introdotto nelle scuole elementari: la cosiddetta *prova del nove*. Si tratta di un metodo di verifica della correttezza del risultato di un calcolo tra numeri interi e consiste, in sostanza, nel calcolare i resti modulo 9 di due numeri, differentemente rappresentati, che si ritiene coincidano. Se questi resti non coincidono, allora sicuramente non coincidono i due numeri. Ad esempio, supponiamo di aver calcolato il prodotto tra due numeri  $a$  e  $b$ , ottenendo  $c$  (discorso analogo varrebbe per una somma anziché un prodotto). Con il metodo della somma delle cifre possiamo calcolare i resti  $a'$  di  $a$ ,  $b'$  di  $b$  e  $c'$  di  $c$  modulo 9. Moltiplichiamo poi tra loro  $a'$  e  $b'$ , e calcoliamo il resto (sempre modulo 9) di  $a'b'$ . Se i nostri calcoli, a partire da  $ab = c$ , sono corretti, allora l'ultimo resto trovato deve essere proprio  $c'$ , altrimenti c'è un errore da qualche parte. Infatti, se  $ab = c$  allora deve essere  $a'b' \equiv_9 ab = c \equiv_9 c'$ . È bene tenere a mente che, viceversa, il fatto che il test "prova del nove" sia passato non garantisce affatto che il calcolo originario sia corretto. Ad esempio, calcoli come  $7 + 2 = 702$  oppure  $33 \cdot 21 = 0$  passano senza problemi la prova del nove.

Veniamo ad un altro criterio di divisibilità molto noto, quello per 11. Ragionando come abbiamo fatto per 9, iniziamo a notare che  $10 \equiv_{11} -1$  e da ciò segue  $10^i \equiv_{11} (-1)^i$ , per ogni intero non negativo  $i$ ; si ha dunque  $10^i \equiv_{11} 1$  se  $i$  è pari e  $10^i \equiv_{11} -1$  se  $i$  è dispari. Allora, per il nostro solito intero positivo  $n = \langle a_t a_{t-1} a_{t-2} \dots a_2 a_1 a_0 \rangle$  abbiamo:

$$n = \sum_{i=0}^t a_i 10^i \equiv_{11} \sum_{i=0}^t (-1)^i a_i = a_0 - a_1 + a_2 - a_3 + \dots + (-1)^t a_t,$$

vale a dire: modulo 11,  $n$  è congruo alla somma delle sue cifre prese a segni alterni, ad iniziare da quella delle unità con segno positivo. Otteniamo così il criterio di divisibilità per 11: *un intero  $n$  è divisibile per 11 se e solo se la differenza tra la somma delle sue cifre di posto pari e quelle di posto dispari è un multiplo di 11*. Meglio, quello che abbiamo dimostrato è che si può calcolare il resto della divisione per 11 di un numero intero positivo calcolando ripetutamente la somma alterna delle cifre menzionata sopra sino a ridursi ad un numero minore di 11; questo numero sarà il resto cercato. Come fatto per i calcoli modulo 9, anche in questo caso, piuttosto che eseguire tutte le somme possiamo usare liberamente le scorciatoie fornite dall'aritmetica modulare per arrivare più rapidamente al risultato. Ad esempio, per lo stesso  $n = 672455978913$  esaminato sopra con riferimento a 9, abbiamo  $n \equiv_{11} 3 - 1 + 9 - 8 + 7 - 9 + 5 - 5 + 4 - 2 + 7 - 6$ , che semplificando in modo ovvio diventa  $n \equiv_{11} 3 - 1 - 8 + 7 + 4 - 2 + 7 - 6 = -1 + 4 + 7 - 6 \equiv_{11} -1 - 6 = -7 \equiv_{11} 4$ . Il resto di  $n$  nella divisione per 11 è dunque 4.

**3.1. Altri criteri.** Serve a qualcosa un criterio di divisibilità per 6? Si potrebbe rispondere di no: un numero intero è divisibile per 6 se e solo se è divisibile sia per 2 che per 3, quindi per sapere se un numero è o meno divisibile per 6 basta applicare i due criteri, già noti, per 2 e per 3. E ci serve davvero un criterio di divisibilità per  $121 = 11^2$ ? Anche in questo caso si potrebbe farne a meno: per stabilire se un certo intero  $n$  è o meno divisibile per 121 si potrebbe applicare ad esso il criterio di divisibilità per 11; se il criterio fallisce allora  $n$  non è divisibile per 121 (per esserlo dovrebbe essere divisibile per 11), se invece  $n$  risulta divisibile per 11 allora si può calcolare  $n' = n/11$  ed applicare lo stesso criterio ad  $n'$ ; infatti  $n$  è divisibile per 121 se e solo se  $n'$  è divisibile per 11. Questo metodo si può estendere a potenze arbitrarie di 11 (o di altri numeri per i quali un criterio di divisibilità sia disponibile); ovviamente non è altrettanto veloce quanto criteri più diretti, come quello per 9.

Queste argomentazioni si possono ulteriormente estendere per raggiungere la conclusione che basti avere criteri di divisibilità per numeri primi per avere criteri di divisibilità per interi arbitrari. Se guardiamo ai numeri interi positivi primi più piccoli, ci accorgiamo di avere a disposizione criteri per 2, 3, 5, 11, ma non per 7, 13, 17 etc.

Esaminiamo il caso di 7. Da un intero positivo  $n$ , isoliamo la cifra  $a$  dell'unità e scriviamo  $n$  come  $10b + a$ ; dunque  $b$  è il numero ottenuto da  $n$  cancellando l'ultima cifra (naturalmente  $b$  ed  $a$  sono il quoziente ed il resto nella divisione di  $n$  per 10). Poiché 21 è multiplo di 7, abbiamo  $20 \equiv_7 -1$ , dunque  $2n = 20b + 2a \equiv_7 -b + 2a$ . Ora, non è difficile riconoscere che  $n$  è divisibile per 7 se e solo se lo è il suo doppio  $2n$ <sup>5</sup>, o, equivalentemente,  $-2n = b - 2a$ . Dunque, per verificare se  $n$  è o meno divisibile

<sup>5</sup>che  $2n$  sia multiplo di 7 se lo è  $n$  è del tutto ovvio. Viceversa, assumiamo che 7 divida  $2n$ . Assumendo nota l'unicità della fattorizzazione degli interi in prodotti di primi, sappiamo che la fattorizzazione di  $2n$  (nella quale deve apparire 7) si ottiene aggiungendo un 2 alla fattorizzazione di  $n$ , quindi 7 deve apparire in questa fattorizzazione, cioè: 7 divide  $n$ .

Più in generale, vale un importantissimo risultato, noto come *lemma di Euclide*: siano  $a, b$  e  $c$  numeri interi. Se  $a$  divide  $bc$  ed è coprimo con  $b$ , allora  $a$  divide  $c$ .

Nell'argomentazione appena svolta abbiamo un po' barato: nella maggior parte delle trattazioni dell'aritmetica il lemma

per 7 possiamo sostituire  $n$  con  $n_1 = b - 2a$ ; questo numero sarà quasi sempre più piccolo di  $n$ : quasi sempre avrà una cifra in meno. Se sappiamo se  $n_1$  è o meno divisibile per 7, a questo punto ci possiamo fermare:  $n$  è divisibile per 7 se e solo se lo è  $n_1$ . Altrimenti, ripetiamo il procedimento partendo da  $n_1$  piuttosto che da  $n$ , e andiamo avanti con la procedura finché non otteniamo la risposta. Un esempio: partiamo da  $n = 314932$ . Dunque, con le notazioni di sopra,  $a = 2$  e  $b = 31493$ ; allora da  $n$  passiamo a  $n_1 = b - 2a = 31489$ . Allo stesso modo, da  $n_1$  passiamo a  $3148 - 2 \cdot 9 = 3130$ , da qui a  $313 - 2 \cdot 0 = 313$ , poi a  $31 - 2 \cdot 3 = 25$ , che sappiamo non essere multiplo di 7. La conclusione è che 7 non divide  $n$ . Si potrebbe velocizzare la procedura e semplificare i conti? Certamente: ragionando come nell'[Esempio 1](#), vediamo che da  $n = 314932$  possiamo “azzerare” il 14; più precisamente  $n \equiv_7 n' := n - 14000 = 300932$ ; inoltre possiamo “ridurre la cifra 9 a 2:  $n' \equiv_7 n'' := n' - 700 = 300232$ . Proseguendo, ed applicando qua e là gli stessi trucchi, da  $n''$  passiamo a  $30023 - 2 \cdot 2 = 30019 \equiv_7 30012$ , da questo a  $3001 - 2 \cdot 2 = 2997 \equiv_7 2220$ , poi a  $222 - 2 \cdot 0 = 222$ , infine a  $22 - 2 \cdot 4 = 18$ , non multiplo di 7.

Osserviamo che, a differenza di quanto accadeva con i criteri di divisibilità esaminati prima, questo criterio di divisibilità per 7, così descritto, non fornisce il resto della divisione del numero per 7, ma si limita solo a stabilire se 7 divide o meno il numero, cioè se questo resto è o non è 0 (in realtà si potrebbe modificare il metodo in modo che si tenga traccia anche del resto, ma al costo di complicarlo).

Possiamo trovare analoghi criteri di divisibilità anche per altri primi. Ad esempio, facciamolo per 13: come nel caso di 7 ci serve un multiplo di 13 che differisca di 1 da un multiplo di 10; andrà bene 39. Definiti  $n$ ,  $a$  e  $b$  come sopra, quindi  $n = 10b+a$ , abbiamo che 13 divide  $n$  se e solo se divide  $4n = 40b+4a \equiv_{13} b+4a$ . Quindi si può decidere se un numero intero è o meno divisibile per 13 con lo stesso metodo usato per 7: l'unica differenza è che, ad ogni passaggio, si sommerà il quadruplo della cifra delle unità anziché sottrarre il doppio. Vediamo direttamente un esempio. Partendo dallo stesso  $n = 314932$  di prima, passiamo a  $31493 + 4 \cdot 2 = 31501$ , poi a  $3150 + 4 \cdot 1 = 3154$ , poi a  $315 + 4 \cdot 4 = 331$ , a  $33 + 4 \cdot 1 = 37$ , che non è multiplo di 13, quindi non lo è  $n$  (volendo avremmo potuto proseguire anche oltre 37, ottenendo  $3 + 4 \cdot 7 = 31$  e quindi  $3 + 4 \cdot 1 = 7$ ). Nulla vieta, anche in questo caso di usare semplificazioni: ad esempio, poiché  $14 \equiv_{13} 1$ , si ha  $n \equiv_{13} 301932$ , quindi saremmo potuti partire da questo numero, o anche da 301802, sfruttando  $93 \equiv_{13} 80$ .

Lo stesso metodo si può applicare, ad esempio, a 17, a 19 ed a primi maggiori. Nel caso di 17 dovremmo usare, con le solite notazioni, la trasformazione  $n \mapsto b - 5a$  (quindi  $-5$  svolge il ruolo che avevano  $-2$  nel caso di 7 e  $4$  nel caso di 13; questo perché  $50 \equiv_{17} -1$ ); per 19 useremo invece  $n \mapsto b + 2a$ , per 23 useremo  $n \mapsto b + 7a$ , per 29 useremo  $n \mapsto b + 3a$ . Chi legge si può divertire a verificare la correttezza di queste affermazioni, a costruirsi esempi, a trovare criteri di divisibilità per altri primi.

#### 4. CALENDARIO DI UN GIRONE ALL'ITALIANA

Supponiamo di volere organizzare un torneo tra un certo numero  $n$  (ovviamente intero e positivo!) di squadre. Il formato prescelto è quello, piuttosto familiare, del *girone all'italiana*: il torneo è articolato in più giornate, in ciascuna delle quali si svolgeranno gli incontri (le partite) tra le squadre; alla fine del torneo ogni squadra dovrà avere incontrato esattamente una volta ciascuna delle altre—the girone di andata (o quello di ritorno) del campionato di calcio di serie A è un esempio di girone all'italiana.

Cerchiamo di essere più precisi. Richiediamo che ogni partita coinvolga esattamente due squadre tra le  $n$  partecipanti al torneo (non è un'osservazione inutile: in un torneo di calcio, basket, rugby, pallavolo lo diamo per scontato, ma in un torneo di bocce o di poker?). Stiamo così anche dicendo che una partita non può essere giocata da una squadra contro se stessa. Altre richieste sono: (1) che ogni giornata coinvolga quante più squadre possibile, ma (2) nessuna squadra abbia due impegni (cioè appaia in due partite) nella stessa giornata. Quante sono, allora, in ciascuna giornata, le partite? Questo è facile: se il numero  $n$  delle squadre è pari, allora ogni giornata consisterà di  $n/2$  partite e vedrà impegnate tutte le squadre; se invece  $n$  è dispari, allora, in ogni giornata, una delle squadre dovrà restare ferma (avrà un *turno di riposo*, come si dice) e le rimanenti  $n - 1$  si affronteranno in  $(n - 1)/2$  partite. E quante sono le giornate in cui si articola il torneo? Beh, qui bisogna esaminare separatamente i casi in cui  $n$  è pari e quello in cui  $n$  è dispari. Siccome ogni squadra deve incontrare (in giornate diverse) una ed una sola volta ciascuna delle altre, ogni squadra dovrà disputare  $n - 1$  partite, e questo è vero in entrambi i casi, ma:

---

di Euclide viene usato per dimostrare il teorema di fattorizzazione unica; quindi la nostra argomentazione assume per noto un teorema più forte di quello che intende provare. Aggiungiamo anche in matematica si preferisce dare un definizione generale di ‘primo’ diversa da quella scolastica, come elemento (ad esempio, numero) per il quale valga una particolare forma del lemma di Euclide.

- se  $n$  è pari ogni squadra gioca precisamente una partita per giornata, quindi il numero delle giornate coincide con quello della partite giocate da una squadra, vale a dire  $n - 1$ ;
- se  $n$  è dispari il calcolo è meno diretto: ciascuna squadra giocherà  $n - 1$  partite (in  $n - 1$  giornate), come per il caso precedente, ma avrà anche dei turni di riposo. Se il numero dei turni di riposo della squadra considerata è  $r$ , il numero delle giornate sarà quindi  $n - 1 + r$ ; questo numero non lo non conosciamo perché non conosciamo ancora  $r$ . Possiamo seguire una strada diversa: contare il numero complessivo di partite: ci sono  $n$  squadre, ciascuna di esse gioca  $n - 1$  partite, abbiamo quindi un totale di  $n(n - 1)$  coppie ordinate  $(S, P)$  costituite da una squadra  $S$  ed una partita  $P$  giocata da  $S$ ; ma ogni partita è giocata da due squadre, quindi per ottenere il numero delle partite dobbiamo dividere il numero di tali coppie per 2. In definitiva, nel torneo verranno giocate in tutto  $n(n - 1)/2$  partite.<sup>6</sup> Poiché ogni giornata consiste di  $(n - 1)/2$  partite, la conclusione è che di giornate ce ne sono  $n$  (il numero  $n(n - 1)/2$  delle partite nel torneo diviso per il numero  $(n - 1)/2$  di partite per giornata). Una conclusione accessoria è anche che ogni squadra osserverà esattamente un turno di riposo, infatti il numero delle giornate del torneo lo avevamo anche calcolato come  $n - 1 + r$ .

Ad esempio, il torneo di rugby delle sei nazioni che, non sorprendentemente, si disputa tra sei squadre (Irlanda, Galles, Scozia, Inghilterra, Francia, Italia) si svolge con un girone all’italiana (di sola andata); il torneo prevede dunque cinque giornate di tre partite ciascuna ( $n = 6$  è il numero delle squadre, che è dunque pari;  $n - 1 = 5$  le giornate;  $n/2$  le partite per giornata). Prima dell’anno 2000 l’Italia non partecipava al torneo, che era ristretto alle altre squadre e si chiamava allora torneo delle cinque nazioni. All’epoca, dunque, il numero delle squadre partecipanti,  $n = 5$ , era dispari, il torneo si svolgeva comunque in cinque giornate, ciascuna delle quali prevedeva due partite ( $2 = (5 - 1)/2$ ) ed una squadra ferma in turno di riposo.

Preparare un calendario per il torneo delle sei nazioni è semplicissimo: quando il numero delle squadre partecipanti è così piccolo, è piuttosto facile stendere un calendario per un girone all’italiana, procedendo per tentativi (si abbinano in qualche modo le squadre per mettere assieme la prima giornata e poi si procede, evitando di ripetere partite già disputate, con le giornate successive), ma quando il numero dei partecipanti sale diventa molto più complicato farlo (procedendo alla cieca si rischia, dopo un certo numero di giornate, di non potere evitare ripetizioni di partite; a quel punto bisogna fare un passo indietro e riprovare), a meno di non disporre di un metodo sistematico da utilizzare allo scopo. Ne discuteremo qui uno molto semplice.

Abbiamo fatto sopra una distinzione tra il caso in cui il numero  $n$  delle squadre partecipanti è pari da quello in cui  $n$  è dispari; effettivamente, nei due casi, i calendari dei tornei hanno aspetto diverso (solo nel secondo caso sono previsti turni di riposo). Potrebbe sembrare necessario cercare due metodi diversi per costruire il nostro calendario; uno da usare nel caso  $n$  sia pari, l’altro nel caso dispari. Fortunatamente non è così: se sappiamo risolvere il nostro problema in uno dei due casi è molto facile risolverlo anche nell’altro.

Supponiamo infatti di saper comporre un calendario per tornei con un arbitrario numero dispari di squadre e trovarci, invece, a doverne comporre uno con un numero pari  $n$  di squadre. Per farlo sarà sufficiente mettere momentaneamente da parte una delle squadre, chiamiamola  $A$ , compilare un calendario per un torneo tra le  $n - 1$  squadre rimanenti (che sono dunque in numero dispari) e poi modificarlo inserendo, ad ogni giornata, una partita tra  $A$  e la squadra che nel primo calendario aveva, in quella giornata, il turno di riposo. Ad esempio, per compilare un calendario per il torneo delle sei nazioni dell’anno 2000 (a sei squadre) sarebbe stato sufficiente prendere il calendario dell’anno precedente (a cinque squadre: mancava l’Italia), che come abbiamo visto prima si svolgeva in cinque giornate di due partite l’una ed aggiungere, in ciascuna giornata, una partita tra l’Italia e la squadra col turno di riposo.

Viceversa, se sappiamo comporre un calendario per tornei con un numero pari di squadre ma abbiamo il compito di organizzare un girone all’italiana per un numero dispari  $n$  di squadre non dovremo fare altro che inventarci una squadra fittizia aggiuntiva, chiamiamola  $F$ , e preparare il calendario per il girone tra le  $n + 1$  squadre (quelle originali più  $F$ ; sono in numero pari quindi siamo capaci di farlo). Per ottenere il calendario desiderato basterà cancellare, ad ogni giornata, la partita che vedeva coinvolta  $F$ ; la squadra avversaria avrà, gioco-forza, turno di riposo. Ad esempio, un calendario per un’edizione del torneo delle cinque nazioni si può ottenere da un calendario per il torneo delle sei nazioni cancellando

---

<sup>6</sup>Abbiamo fatto del semplice calcolo combinatorio, chi ne ricorda un po’ riconosce qui il coefficiente binomiale  $\binom{n}{2} = n(n - 1)/2$ , che è il numero dei sottoinsiemi costituiti da due elementi in un insieme costituito da  $n$  elementi: osservare che una partita si può anche riguardare come la selezione di un sottoinsieme di due elementi nell’insieme di tutte le squadre.

sistematicamente le partite dell'Italia e lasciando a riposo, giornata per giornata, l'avversario di turno dell'Italia.

Non è difficile verificare (e chi legge è invitato a farlo) che i calendari ottenuti da questi due procedimenti verificano le richieste fatte (che ogni squadra incontri ciascuna altra squadra esattamente una volta nel torneo, eccetera).

A questo punto sappiamo che per descrivere un metodo generale per la stesura di un calendario per un girone all'italiana basta limitarsi a farlo nel caso in cui il numero delle squadre è pari, o, in alternativa, nel caso in cui questo numero è dispari. Anche se la prima opzione è quella che corrisponde al caso più consueto (in genere i tornei sono organizzati tra un numero pari di squadre) ci pare più semplice, dal punto di vista matematico, la descrizione della procedura nel secondo caso, quindi supporremo che il numero  $n$  delle squadre partecipanti al torneo sia dispari.

I dati di partenza sono dunque questi: un intero positivo dispari  $n$  e  $n$  squadre (distinte tra loro), che indichiamo come  $S_1, S_2, \dots, S_n$ . In accordo con i calcoli svolti sopra, dobbiamo definire  $n$  giornate del nostro torneo, che possiamo ovviamente chiamare prima giornata, seconda giornata, e così via, sino alla  $n$ -esima giornata. Ciascuna delle giornate è specificata dalle  $(n-1)/2$  partite che vi si svolgono (tra  $n-1$  delle squadre; la rimanente ha il turno di riposo). La definizione che proponiamo è questa:

*scelti comunque gli interi  $i, j, k$  nell'insieme  $\{1, 2, 3, \dots, n\}$ , nella  $k$ -giornata si svolge la partita tra le squadre  $S_i$  e  $S_j$  se e solo se  $i \neq j$  e  $i + j \equiv_n k$ .*

Ad esempio, se  $n = 15$ , nella prima giornata giocheranno tra loro  $S_4$  ed  $S_{12}$ , perché  $4 + 12 \equiv_{15} 1$ ; nella tredicesima, invece,  $S_4$  giocherà contro  $S_9$ , perché  $4 + 9 \equiv_{15} 13$ .

Verifichiamo che la definizione che abbiamo proposto fornisce un calendario “corretto”. Si tratta di controllare che il calendario che stiamo definendo soddisfa tutti i vincoli che avevamo richiesto. Innanzitutto, nessuna squadra gioca mai con se stessa: la nostra definizione impone che affinché si possa svolgere una partita tra le squadre  $S_i$  e  $S_j$  si debba avere  $i \neq j$ . Dunque, ogni partita coinvolge precisamente due squadre (distinte). È vero che ogni squadra incontra ciascuna altra squadra esattamente una volta nel torneo? Sì, perché se  $S_i$  e  $S_j$  sono due tra le nostre squadre, e  $i \neq j$ , sappiamo che esiste uno ed un solo  $k_{ij} \in \{1, 2, 3, \dots, n\}$  tale che  $i + j \equiv_n k_{ij}$  (detto  $r$  il resto nella divisione di  $i + j$  per  $n$ , se  $r \neq 0$ , allora questo  $k_{ij}$  è proprio  $r$ , altrimenti  $k_{ij}$  è  $n$ ). Questo significa che  $S_i$  e  $S_j$  si incontreranno esattamente una volta: nella  $k_{ij}$ -esima giornata. Notiamo anche che quando  $S_i$  gioca con  $S_j$ , si ha anche che  $S_j$  gioca con  $S_i$  (meno male! Se così non fosse le “partite” non sarebbero ben definite), questo è conseguenza della proprietà commutativa dell’addizione. Quante sono le partite che si svolgono in una fissata ora una giornata, diciamo la  $k$ -esima? Fissato  $i \in \{1, 2, \dots, n\}$ , esiste uno ed un solo  $j \in \{1, 2, \dots, n\}$  tale che  $i + j \equiv_n k$ : è quel’unico  $j$  nell’insieme dato tale che  $j \equiv_n k - i$ . La nostra definizione assicura che nella  $k$ -esima giornata  $S_i$  non potrà giocare con alcuna squadra diversa da  $S_j$ ; quindi  $S_i$  gioca al massimo una partita in questa giornata. Ora si danno due possibilità: o  $i \neq j$ , ed allora  $S_i$  gioca con  $S_j$ , oppure  $i = j$  e quindi  $S_i$  non gioca con  $S_j$ ; ma in questo secondo caso  $S_i$ , nella  $k$ -esima giornata non gioca proprio ed osserva un turno di riposo. Benissimo, quante sono le squadre che riposano nella  $k$ -esima giornata? Da quanto abbiamo appena visto segue subito che una squadra  $S_i$  riposa nella  $k$ -esima giornata se e solo se si ha  $2i \equiv_n k$ , quindi bisogna stabilire quanti sono gli  $i \in \{1, 2, \dots, n\}$  tali che  $2i \equiv_n k$ . Lo si potrebbe fare immediatamente utilizzando la teoria delle equazioni congruenziali (vedi oltre per qualche cenno a riguardo), ma possiamo anche arrivarci per via diretta. Se  $k$  è pari, un tale  $i$  è certamente  $i/2$ . Se  $k$  è dispari, invece,  $n+k$  è pari (perché  $n$ , ricordiamo, è dispari), dunque  $(n+k)/2$  è un intero; ponendo appunto  $i = (n+k)/2$  abbiamo di nuovo  $2i = n+k \equiv_n k$ . Dunque, in ciascun caso almeno un  $i \in \{1, 2, \dots, n\}$  verifica la condizione.<sup>7</sup> È possibile che più di una squadra riposi nella  $k$ -esima giornata? Vediamo: supponiamo che riposino sia  $S_i$  che  $S_{i'}$ , dove, naturalmente,  $i, i' \in \{1, 2, \dots, n\}$ . Allora sia  $2i$  che  $2i'$  sono congrui a  $k$  modulo  $n$ ; in particolare  $2i \equiv_k 2i'$ , vale a dire:  $k$  divide  $2(i - i')$ . Poiché  $k$  è dispari, se ne ricava che  $k$  divide  $i - i'$ , ovvero  $i' \equiv_n i$ .<sup>8</sup> Ricordando che gli elementi di  $\{1, 2, \dots, n\}$  sono a due a due non congrui modulo  $n$ , otteniamo  $i' = i$ , cioè  $S_i = S_{i'}$ . In questo modo abbiamo provato che

<sup>7</sup>in altri termini: almeno una squadra riposa nella  $k$ -esima giornata. Del resto questo segue anche dal fatto che le partite che si svolgono nella giornata coinvolgono un numero totale pari di squadre (due per partita, tutte diverse tra loro); essendo il numero delle squadre partecipanti dispari, è chiaro che almeno una di esse non gioca.

<sup>8</sup>la cosa è abbastanza intuitiva, ma dipende in ultima analisi dal lemma di Euclide; vedi la nota 5.

solo una squadra riposa nella  $k$ -esima giornata. A questo punto siamo certi del fatto che il nostro metodo fornisce un calendario corretto per un girone all'italiana.

Possiamo descrivere questo metodo in modo ancora più sintetico facendo riferimento alle tavole di Cayley, che abbiamo discusso [in precedenza](#): il calendario che abbiamo descritto è codificato nella tavola di Cayley per l'addizione in  $\mathbb{Z}_n$ . Lo possiamo comprendere bene con un esempio concreto. Poniamo  $n = 7$  e scriviamo la tavola di Cayley per l'addizione in  $\mathbb{Z}_7$ . Per poter mantenere le stesse notazioni che stiamo usando in questa sezione, utilizziamo però 7 anziché 0 per rappresentare la classe  $[0]_7$ :

$+$	7	1	2	3	4	5	6
7	7	1	2	3	4	5	6
1	1	2	3	4	5	6	7
2	2	3	4	5	6	7	1
3	3	4	5	6	7	1	2
4	4	5	6	7	1	2	3
5	5	6	7	1	2	3	4
6	6	7	1	2	3	4	5

Leggiamo in questa tabella il calendario: assumendo  $i \neq j$ , la squadra  $S_i$  incontra la squadra  $S_j$  nella giornata indicata dal numero nella riga intestata da  $i$  e colonna intestata da  $j$ , dal momento che questo numero è congruo a  $i + j$ . Ad esempio,  $S_2$  incontra  $S_6$  nella prima giornata, ed incontra  $S_3$  nella quinta. Quindi i numeri che leggiamo nella tabella, al di fuori della diagonale da “alto-sinistra” a “basso-destra” ci dicono in quale giornata si incontreranno le squadre che descrivono la posizione del numero. E cosa ci dicono invece i numeri sulla diagonale? Beh, un numero che appare sulla diagonale, alla riga intestata da un certo  $i$ , è anche nella colonna intestata da  $i$ , quindi rappresenta la classe  $[i]_7 + [i]_7 = [2i]_7$ ; in altre parole è congruo a  $2i$ . Di conseguenza quel numero indica la giornata in cui la squadra  $S_i$  ha il suo turno di riposo; ad esempio,  $S_5$  riposa nella terza giornata,  $S_6$  nella quinta.

Abbiamo mostrato come, per un qualsiasi intero positivo dispari  $n$  si possa usare l'addizione in  $\mathbb{Z}_n$  per organizzare un girone all'italiana per un torneo con  $n$  squadre. Lo stesso procedimento si può adottare per ottenere altri calendari, utilizzando altri tipi di strutture al posto di  $\mathbb{Z}_n$ . Forse chi legge sa cosa sia un gruppo abeliano (è un tipo di struttura algebrica di cui  $\mathbb{Z}_n$ , con l'operazione di addizione, è un esempio; ‘abeliano’ è solo un modo complicato per dire che vale la proprietà commutativa); se lo sa potrà facilmente riconoscere se può usare la tavola di Cayley di un qualsiasi gruppo abeliano con  $n$  elementi per costruire, con lo stesso metodo che abbiamo usato qui, un calendario per un girone all'italiana tra  $n$  squadre. Questo dipende dal fatto che le tavole di Cayley dei gruppi abeliani hanno le due proprietà che abbiamo sfruttato per costruire i nostri calendari a partire dall'addizione in  $\mathbb{Z}_n$ : sono simmetriche rispetto alla diagonale (perché l'operazione è commutativa) e, come nel Sudoku, non presentano mai ripetizioni dello stesso simbolo in alcuna riga. Questa seconda proprietà (che si chiama cancellabilità) avrà un ruolo importante [nella prossima sezione](#).

## 5. INVERTIBILI, FERMAT ED EULERO

Torniamo ora a qualcosa di carattere più teorico (e quindi, a giudizio di chi scrive, probabilmente più interessante). Nello studio degli anelli  $\mathbb{Z}_m$  ha grande importanza la descrizione degli elementi invertibili. Abbiamo già incontrato questa nozione: un elemento  $[a]_m$  di  $\mathbb{Z}_m$  (consideriamo fissato l'intero positivo  $m$ ) è invertibile se e solo se esiste un  $[b]_m \in \mathbb{Z}_m$  tale che  $[a]_m[b]_m = [1]_m$ . Si dimostra (lo si potrebbe far seguire dalla proprietà di cancellabilità che vedremo tra poco) che se  $[a]_m$  è invertibile allora esiste esattamente un  $[b]_m$  con la proprietà richiesta; questo  $[b]_m$  si chiama l'*inverso* di  $[a]_m$ .<sup>9</sup>

Notiamo che  $[a]_m[b]_m$  è di per sé  $[ab]_m$ ; la condizione di invertibilità si può dunque riformulare così:  $[a]_m$  è invertibile se e solo se esiste  $b \in \mathbb{Z}$  tale che  $[ab]_m = [1]_m$ , ovvero  $ab \equiv_m 1$ . Ad esempio, abbiamo già visto che in  $\mathbb{Z}_5$  sono invertibili tutti gli elementi tranne  $[0]_5$ , mentre in  $\mathbb{Z}_6$  sono invertibili solo  $[1]_6$

<sup>9</sup>la definizione di elemento invertibile e di inverso si dà in contesti molto più generali, quella che abbiamo formulato qui ne è giusto un caso particolare. Quando la si dà in altre strutture, il ruolo che qui svolge  $[1]_m$  è riservato all'elemento neutro della struttura (anche dell'elemento neutro si può dimostrare l'unicità). Aggiungiamo anche, giusto per scrupolo di precisione, che la definizione va modificata (è un po' più elaborata) nel caso in cui la struttura considerata non sia commutativa, come invece è quella che stiamo considerando in  $\mathbb{Z}_m$ .

e  $[-1]_6 = [5]_6$  (ciascuno dei quali coincide col suo inverso); in  $\mathbb{Z}_{10}$  è, tra gli altri, invertibile  $[3]_{10}$ , il cui inverso è  $[7]_{10}$  (verificare!).

Esiste un semplice modo per stabilire se un elemento di  $\mathbb{Z}_m$  è o meno invertibile. Per descriverlo possiamo fare riferimento alle cosiddette *equazioni congruenziali* (di primo grado). Cosa sono? Essenzialmente ordinarie equazioni considerate in  $\mathbb{Z}_m$  anziché in uno degli abituali insiemi di numeri. Una equazione di primo grado in  $\mathbb{Z}_m$ , in una incognita  $X$ , è una equazione della forma

$$AX = B \tag{*}$$

dove  $A$  e  $B$  sono due elementi di  $\mathbb{Z}_m$ , cioè due classi di resto modulo  $m$ . Ad esempio,  $[2]_7X = [3]_7$  è un'equazione del genere, per  $m = 7$ . Naturalmente risolvere l'equazione (\*) significa trovare una classe (o le classi) che sostituite ad  $X$  rendano vera l'uguaglianza. Siano  $a$  un numero in  $A$  e  $b$  un numero in  $B$ , quindi  $A = [a]_m$  e  $B = [b]_m$ . Siccome  $X$  rappresenta una classe di resto, possiamo anche scrivere  $X = [x]_m$ , dove  $x$  è una incognita che rappresenta un numero intero. Allora l'equazione (\*) si può anche scrivere come  $[a]_m[x]_m = [b]_m$ , ovvero:

$$ax \equiv_m b. \tag{**}$$

Questa si chiama una equazione congruenziale; è, meglio ripeterlo, una forma equivalente di (\*). Ad esempio,  $[2]_7X = [3]_7$  diventa l'equazione congruenziale  $2x \equiv_7 3$ , della quale 5 è una soluzione. Non sempre un'equazione congruenziale come (\*\*) ha soluzione; esiste un importante teorema che fornisce un criterio affinché ne abbia. Lo enunciamo ma non lo dimostriamo:

**Teorema** (Criterio di esistenza di soluzioni per equazioni congruenziali). *Scelti comunque  $a, b \in \mathbb{Z}$  e l'intero positivo  $m$ , detto  $d$  un massimo comun divisore tra  $a$  e  $m$ , l'equazione congruenziale (\*\*) ha soluzione se e solo se  $d$  divide  $b$ .*

Ad esempio,  $12x \equiv_{58} 73$  non ha soluzioni, mentre  $12x \equiv_{58} 74$  ne ha, perché il massimo comun divisore positivo tra 12 e 58 è 2, che divide 74 ma non 73. Esiste un metodo esplicito per trovare tutte le (eventuali) soluzioni di una equazione congruenziale di primo grado, ma non ci addentriamo nella sua descrizione. Dimostriamo invece la conseguenza che ci interessa del teorema appena enunciato:

**Corollario.** *Siano  $a$  un intero e  $m$  un intero positivo. Allora  $[a]_m$  è un elemento invertibile di  $\mathbb{Z}_m$  se e solo se  $a$  ed  $m$  sono coprimi.*

*Dimostrazione.* Dire che  $[a]_m$  è invertibile significa dire che esiste  $b \in \mathbb{Z}$  tale che  $ab \equiv_m 1$ , ovvero: che l'equazione congruenziale  $ax \equiv_m 1$  ha soluzione. Se  $d$  è il massimo comun divisore positivo tra  $a$  e  $m$ , allora, per il criterio di esistenza di soluzioni per equazioni congruenziali, appena enunciato,  $ax \equiv_m 1$  ha soluzione se e solo se  $d$  divide 1. Ma l'unico divisore positivo di 1 è 1 stesso, quindi questa condizione equivale a richiedere che  $d$  sia 1, cioè che  $a$  ed  $m$  siano coprimi. Dunque,  $[a]_m$  è invertibile se e solo se  $a$  ed  $m$  sono coprimi.  $\square$

Questo risultato spiega la differenza che avevamo osservato tra  $\mathbb{Z}_5$  e  $\mathbb{Z}_6$  quando avevamo costruito le [tavole di Cayley](#) delle rispettive operazioni di moltiplicazione. Il numero 5 è primo, da questo si deduce che ogni numero intero compreso tra 1 e 4 è coprimo con 5; quindi tutti gli elementi di  $\mathbb{Z}_5$  escluso  $[0]_5$  sono invertibili (che  $[0]_m$  non possa mai essere invertibile, per alcun  $m > 1$ , è molto facile da riconoscere, anche senza far ricorso al [corollario](#) appena provato). Invece la situazione è diversa per  $\mathbb{Z}_6$ : nessuno tra 2, 3 e 4 è coprimo con 6, quindi  $[2]_6, [3]_6$  e  $[4]_6$  non sono invertibili. Possiamo generalizzare il discorso fatto per 5: se  $m$  è un numero primo e  $a$  è un intero tale che  $1 \leq a < m$ , allora  $a$  ed  $m$  sono certamente coprimi, quindi  $[a]_m$  è invertibile (se  $a$  non fosse coprimo con  $m$  dovrebbe essere divisibile per un primo divisore anche di  $m$ . Ma l'unico primo che divida  $m$  è  $m$  stesso, quindi  $a$ , per non essere coprimo con  $m$ , dovrebbe essere un multiplo di  $m$ , cosa impossibile dal momento che  $0 < a < m$ ). Dunque, se  $m$  è primo ogni elemento di  $\mathbb{Z}_m$  diverso da  $[0]_m$  è invertibile; come detto in precedenza, si esprime questo fatto dicendo che  $\mathbb{Z}_m$  è, in questo caso, un campo. Si potrebbe anche dimostrare (e non è difficile, chi legge è invitato a provarci) che vale anche il viceversa: se  $\mathbb{Z}_m$  è un campo allora  $m$  è primo.

**5.1. La funzione di Eulero.** Il numero degli elementi invertibili in  $\mathbb{Z}_m$  è espresso da una funzione (dall'insieme dei numeri interi positivi in sé), tradizionalmente indicata con  $\varphi$  e nota come *funzione di Eulero*.<sup>10</sup> Dunque, per ogni intero positivo  $m$ , con  $\varphi(m)$  si intende il numero degli elementi invertibili in  $\mathbb{Z}_m$ . Dalla [descrizione degli elementi di  \$\mathbb{Z}\_m\$](#)  e dal [corollario](#) che abbiamo dimostrato poco fa segue che  $\varphi(m)$  è anche il numero degli interi  $a$  coprimi con  $m$  e tali che  $0 \leq a < m$ . Ad esempio, se  $m = 2$ , l'unico

<sup>10</sup>vengono usati anche altri nomi: funzione (o indicatore) di Eulero-Gauss, funzione totiente.

intero con la proprietà richiesta per  $a$  è 1, quindi  $\varphi(2) = 1$ , se  $m = 3$  abbiamo invece due tali interi (1 e 2) e lo stesso vale se  $m = 4$  (in questo caso gli interi sono 1 e 3). In un certo modo abbiamo già visto che  $\varphi(5) = 4$  e  $\varphi(6) = 2$ .

In alcuni casi il calcolo di  $\varphi(m)$  è molto semplice. Se  $m$  è primo, infatti, sappiamo che tutti gli elementi di  $\mathbb{Z}_m$  tranne uno di essi ( $[0]_m$ ) sono invertibili; ma allora, siccome  $\mathbb{Z}_m$  ha precisamente  $m$  elementi,  $\varphi(m) = m - 1$ . Più in generale, cosa succede se  $m$  è una potenza di primo, poniamo  $m = p^n$ , dove  $p$  è primo e  $n$  è un intero positivo? Duplicando un ragionamento svolto sopra, vediamo che un numero intero  $a$ , per non essere coprimo con  $p^n$ , deve essere divisibile per un qualche primo che divida anche  $p^n$ , ma l'unico primo che divida  $p^n$  è  $p$ . Allora, per un arbitrario  $a \in \mathbb{Z}$ , i casi possibili sono due: o  $p$  divide  $a$  (ed allora  $a$  non è coprimo con  $p^n$ ), oppure  $p$  non divide  $a$  (ed allora  $a$  è coprimo con  $p^n$ ). Di conseguenza,  $\varphi(p^n)$  è uguale al numero degli interi compresi tra 0 e  $p^n - 1$  che non siano multipli di  $p$ . Siccome i numeri compresi tra 0 e  $p^n - 1$  sono in tutto  $p^n$  e, tra essi i multipli di  $p$  sono quelli della forma  $kp$ , al variare di  $k$  tra 0 e  $p^{n-1} - 1$ , che sono  $p^{n-1}$ , concludiamo che

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

Una proprietà estremamente significativa della funzione di Eulero, che si chiama moltiplicatività, è:

$$\text{se } a \text{ e } b \text{ sono due numeri interi positivi tra loro coprimi, allora } \varphi(ab) = \varphi(a)\varphi(b).$$

Non dimostriamo questa proprietà, diciamo però (così abbiamo la scusa per imparare qualcosa in più) che dipende da un teorema noto come *teorema cinese dei resti*: se  $a$  e  $b$  sono due numeri interi tra loro coprimi e  $r, s \in \mathbb{Z}$ , allora esiste un intero  $n$  tale che  $n \equiv_a r$  e  $n \equiv_b s$ ; l'insieme di tali interi  $n$  costituisce una classe di resto modulo  $ab$ .

La moltiplicatività, assieme alle osservazioni svolte prima, permette di calcolare i valori della funzione di Eulero per ogni intero positivo  $n$  di cui conosciamo la decomposizione in prodotto di primi. Supponiamo infatti  $n = p_1^{\lambda_1} p_2^{\lambda_2} \cdots p_t^{\lambda_t}$ , dove  $t$  è un intero non negativo, i  $p_i$  sono numeri primi (positivi) a due a due distinti e ciascuno degli esponenti  $\lambda_i$  è un intero positivo. Poiché ciascuno dei fattori  $p_i^{\lambda_i}$  è coprimo col prodotto degli altri fattori di  $n$ , possiamo usare ripetutamente la moltiplicatività della funzione di Eulero per ottenere:

$$\varphi(n) = \varphi(p_1^{\lambda_1} p_2^{\lambda_2} \cdots p_t^{\lambda_t}) = \varphi(p_1^{\lambda_1})\varphi(p_2^{\lambda_2}) \cdots \varphi(p_t^{\lambda_t}) = (p_1^{\lambda_1} - p_1^{\lambda_1-1})(p_2^{\lambda_2} - p_2^{\lambda_2-1}) \cdots (p_t^{\lambda_t} - p_t^{\lambda_t-1}).$$

Ad esempio,  $\varphi(10) = \varphi(2)\varphi(5) = (2 - 1)(5 - 1) = 4$ , infatti si verifica direttamente che le classi di resto invertibili in  $\mathbb{Z}_{10}$  sono quattro:  $[1]_{10}$ ,  $[3]_{10}$ ,  $[-3]_{10} = [7]_{10}$  e  $[-1]_{10} = [9]_{10}$ . Un altro esempio: poiché  $6000 = 2^4 \cdot 3 \cdot 5^3$ , si ha  $\varphi(6000) = (2^4 - 2^3) \cdot (3 - 1) \cdot (5^3 - 5^2) = 8 \cdot 2 \cdot 100 = 1600$ . Osserviamo anche che  $\varphi(6000)$  non è uguale a  $\varphi(6)\varphi(1000) = 800$  (verificarlo!), benché  $6000 = 6 \cdot 1000$ : per poter utilizzare la proprietà di moltiplicatività della funzione di Eulero bisogna che i fattori del prodotto che si considera siano coprimi, ma 6 e 1000 non lo sono.

## 5.2. Il teorema di Fermat-Eulero.

Il teorema è questo:

**Teorema** (Fermat-Eulero). *Siano  $m$  un intero positivo e  $a$  un intero coprimo con  $m$ . Allora  $a^{\varphi(m)} \equiv_m 1$ .*

Questo teorema fu dimostrato da Leonhard Euler (svizzero, uno dei grandissimi nella storia della matematica) nel 1763, estendendo un precedente teorema, annunciato da Pierre de Fermat (un altro grande matematico, questa volta francese, che non aveva l'abitudine di rendere pubbliche le dimostrazioni dei suoi teoremi) in una lettera ad un amico nel 1640.

**Teorema** (Piccolo Teorema di Fermat). *Siano  $p$  un numero primo ed  $a$  un intero. Allora  $a^p \equiv_p a$ .*

Questo secondo teorema si può dimostrare molto facilmente come conseguenza del precedente (che però è storicamente successivo). Infatti, come ormai abbiamo visto in un paio di occasioni, dal momento che  $p$  è primo solo due casi sono possibili: o  $p$  divide  $a$  oppure  $a$  e  $p$  sono coprimi. Nel primo caso  $p$  divide anche  $a^p$ , ovviamente, quindi  $a^p \equiv_p 0 \equiv_p a$ , nel secondo, applicando il teorema di Fermat-Eulero, siccome

$\varphi(p) = p - 1$  abbiamo  $a^{p-1} = a^{\varphi(p)} \equiv_p 1$ , e quindi (moltiplicando per  $a$ )  $a^p = a \cdot a^{p-1} \equiv_p a \cdot 1 = a$ . In entrambi i casi, dunque,  $a^p \equiv_p a$ , come richiesto dall'enunciato.<sup>11</sup>

Vediamo qualche esempio: abbiamo fatto dei calcoli per la determinazione di giorni della settimana, quindi in aritmetica modulo 7. Bene, ora sappiamo che, ogni numero  $a$  che non sia multiplo di 7 verifica  $a^6 \equiv_7 1$ , per il teorema di Fermat-Eulero; saperlo da prima ci avrebbe risparmiato qualche calcolo. Ma anche, ad esempio, scopriamo, senza dover fare conti, che  $379^{1600} - 1$  è un multiplo di 6000; infatti abbiamo visto che  $\varphi(6000) = 1600$  e, inoltre, 379, non essendo divisibile né per 2, né per 3 né per 5, è coprimo con 6000 e quindi  $379^{1600} \equiv_{6000} 1$ . Esempi del genere dovrebbero convincere chi legge del gran significato e della grande utilità del teorema di Fermat-Eulero.

Veniamo ad una sua dimostrazione. Ne possiamo fornire una del tutto elementare, che parte da due semplici considerazioni sugli elementi invertibili degli anelli  $\mathbb{Z}_m$  (e restano valide anche in contesti molto più generali).

La prima: il prodotto tra due elementi invertibili è ancora invertibile. Infatti, se  $A$  e  $B$  sono due classi di resto invertibili in  $\mathbb{Z}_m$ , con inversi, rispettivamente,  $A'$  e  $B'$ , allora si ha  $(AB)(A'B') = A(BB')A' = A[1]_mA' = AA' = [1]_m$ . Questo, (riguardarsi le definizioni [all'inizio di questa sezione](#)) significa che  $AB$  è invertibile, come volevamo provare, ed anche che  $B'A'$  è il suo inverso.

La seconda: ogni elemento invertibile  $A$  verifica questa proprietà, detta *cancellabilità*:

$$\text{per ogni } X, Y \in \mathbb{Z}_m \text{ se } AX = AY \text{ allora } X = Y.$$

Infatti, se  $AX = AY$ , moltiplicando per l'inverso  $A'$  di  $A$  otteniamo  $A'AX = A'AY$ , quindi  $X = A'AX = A'AY = Y$ .<sup>12</sup>

Una parentesi: avevamo già accennato alla cancellabilità nella sezione precedente, in relazione ai calendari dei gironi all'italiana ed alla tavole di Cayley. Ribadiamo questo punto: la cancellabilità di un elemento  $A$  di  $\mathbb{Z}_m$  significa precisamente che, nella tavola di Cayley di  $\mathbb{Z}_m$  rispetto alla moltiplicazione, la riga di  $A$  non presenta ripetizioni. Infatti, dire che ci sono ripetizioni significa dire che, in due posizioni distinte della riga, quindi in corrispondenza di due colonne distinte, quella di un elemento  $X$  e quella di un elemento  $Y$ , con  $X \neq Y$ , dobbiamo avere lo stesso elemento  $Z$ , quindi  $Z = AX = AY$ . Questo è precisamente ciò che la proprietà di cancellabilità per  $A$  esclude.

Torniamo alla dimostrazione. Elenchiamo gli elementi invertibili di  $\mathbb{Z}_m$ :

$$X_1, X_2, X_3, \dots, X_{\varphi(m)}$$

(ricordiamo che gli invertibili di  $\mathbb{Z}_m$  sono in tutto  $\varphi(m)$ , di conseguenza questa è una lista priva di ripetizioni: le classi di resto elencate sono a due a due distinte). Sia  $A$  uno di essi. Se moltiplichiamo ciascun elemento della lista precedente per  $A$  otteniamo ancora una lista:

$$AX_1, AX_2, AX_3, \dots, AX_{\varphi(m)}$$

di elementi invertibili di  $\mathbb{Z}_m$  (perché il prodotto di due elementi invertibili è sempre invertibile). Inoltre questi elementi sono a due a due distinti: se così non fosse ci sarebbero un  $i$  ed un  $j$  in  $\{1, 2, 3, \dots, \varphi(m)\}$  tali che  $i \neq j$  e  $AX_i = AX_j$ , ma allora  $X_i = X_j$  per la cancellabilità di  $A$ ; questo è impossibile perché la nostra prima lista non presenta ripetizioni. Dunque, anche la seconda lista consiste di  $\varphi(m)$  elementi invertibili di  $\mathbb{Z}_m$ . Siccome  $\mathbb{Z}_m$  ha esattamente  $\varphi(m)$  elementi, anche questa seconda lista comprende tutti (e soli) gli elementi invertibili di  $\mathbb{Z}_m$ , senza ripetizioni. In conclusione: le due liste elencano esattamente gli stessi elementi; esse possono differire solo per l'ordine in cui questi appaiono. Moltiplichiamo tra loro le classi nella prima lista da una parte, e quelle nella seconda lista dall'altra. Siccome vale la proprietà commutativa, quindi l'ordine dei fattori è irrilevante ai fini del calcolo del prodotto, otteniamo nei due

<sup>11</sup>Molto spesso la presentazione dei risultati della matematica riflette la loro storia in modo poco fedele. Anche questo è il caso: andando a rileggere la lettera originale di Fermat si scopre che Fermat non enunciò il teorema nella forma che gli abbiamo dato, che è quella diventata canonica ed usata dai matematici moderni: un enunciato valido per ogni intero  $a$ . Fermat, invece, si limitò ad enunciare il teorema solo nel caso in cui  $p$  non divida  $a$  (che è comunque l'unico caso non banale del teorema). Quindi quello che Fermat enunciò è proprio il teorema che abbiamo chiamato di Fermat-Eulero ristretto, nelle ipotesi, al caso in cui  $m$  sia primo.

Questo non è un caso isolato: quasi mai succede, in matematica, che il ‘Teorema di Tizio’ sia davvero ciò che Tizio aveva detto. Ciò dipende, probabilmente, non da incuria o sciatteria, ma proprio dalla natura della matematica: una disciplina in cui ciò che è stato scoperto una volta varrà per sempre, ma verrà anche rielaborato e riformulato senza sosta da chi lo guarda da prospettive sempre diverse.

<sup>12</sup>un’osservazione di carattere algebrico: qui è essenziale la proprietà associativa.

casi lo stesso risultato, chiamiamolo  $Y$ :

$$\prod_{i=1}^{\varphi(m)} X_i = Y = \prod_{i=1}^{\varphi(m)} (AX_i).$$

Raggruppando i fattori indicati come  $A$  nel secondo prodotto otteniamo anche

$$Y = \prod_{i=1}^{\varphi(m)} (AX_i) = \left( \prod_{i=1}^{\varphi(m)} A \right) \cdot \left( \prod_{i=1}^{\varphi(m)} X_i \right) = A^{\varphi(m)} Y.$$

Dunque  $[1]_m Y = Y = A^{\varphi(m)} Y$ . Ricordando che  $Y$  è invertibile (in quanto prodotto di invertibili) e quindi cancellabile, ne deduciamo  $[1]_m = A^{\varphi(m)}$ .

Abbiamo così verificato che la potenza  $\varphi(m)$ -esima di un qualsiasi elemento invertibile di  $\mathbb{Z}_m$  è la classe  $[1]_m$ . Con questo siamo quasi alla conclusione. Sia infatti  $a$  un intero coprimo con  $m$  e poniamo  $A = [a]_m$ . Sappiamo che  $A$  è invertibile (lo avevamo visto con il [corollario](#) nella prima parte di questa sezione), quindi  $A^{\varphi(m)} = [1]_m$ . Ma  $A^{\varphi(m)} = ([a]_m)^{\varphi(m)} = [a^{\varphi(m)}]_m$ , dunque  $[a^{\varphi(m)}]_m = [1]_m$ , ovvero

$$a^{\varphi(m)} \equiv_m 1,$$

che è precisamente ciò che volevamo provare. La dimostrazione del teorema di Fermat-Eulero è così completa.

## 6. PER APPROFONDIRE . . .

La lettura di questo articolo può essere utilmente complementata da quella di un articolo di natura analoga: M.R. Celentani, *Aritmetica modulare: una proposta didattica*, Periodico di Matematiche, vol. 6 serie XI, anno CXXIV, pag. 19.

Chi ha trovato interessante la matematica discussa in questi articoli, può approfondire l'argomento consultando una fonte che ne fornisca una trattazione più sistematica, come ad esempio un testo universitario di algebra.

Uno dei motivi dell'inclusione del teorema di Fermat-Eulero in queste note è che questo teorema trova applicazione in crittografia: esso è infatti alla base del primo protocollo crittografico moderno (a chiave pubblica), il protocollo RSA. Si è preferito però evitare di trattare qui questo genere di applicazioni; presentazioni della crittografia, a qualsiasi livello, ce ne sono a bizzeffe. Una che, come questo articolo, è rivolta a studenti delle scuole superiori è in un mio articolo divulgativo (*Matematica e crittografia*) reperibile all'indirizzo <http://www.dma.unina.it/~cutolo/didattica/varia/coinor-cutolo.pdf>.

# Aritmetica Modulare

$$m = 9 \quad 7284512 \quad \sum_{i=0}^t a_i \cdot 10^i \rightarrow \equiv_9 \sum_{i=0}^t a_i$$

" $a_t a_{t-1} \dots a_2 a_1 a_0$ " =  $\sum_{i=0}^t a_i \cdot 10^i$

$10 \equiv_9 1 \quad \forall i \in \mathbb{N}$     $10^i = \underbrace{10 \cdot 10 \dots \cdot 10}_i \equiv_9 1 \cdot 1 \cdot \dots \cdot 1 = 1$

$\equiv_9 (-1)^i$

"Oggi è venerdì. Tra  $2^{104}$  giorni, che giorno sarà?"

$$2^3 = 8 \equiv_7 1 \rightarrow 2^{3k} \equiv_7 1^k = 1 \rightarrow 2^{3k+1} \equiv_7 1 \cdot 2^1 = 2^1$$

$$104 \bmod 3 = 2 \quad 104 \equiv_3 5 \equiv 2 \rightarrow \exists k \in \mathbb{N} \quad 104 = 3k + 2$$

$$2^{104} \equiv_7 2^2 = 4$$

Quindi: tra  $2^{104}$  sarà martedì.



## Anelli e mod

$$\mathbb{Z}_2 = \left\{ [0]_2, [1]_2 \right\} \text{ campo, anello booleano } (\mathbb{Z}, +, \cdot)$$

$$+ \leftrightarrow \text{XOR} \leftrightarrow \Delta$$

$$\cdot \leftrightarrow \text{AND} \leftrightarrow \cap$$

$\left( \begin{array}{l} \text{si rimanda all'ultima pagina di} \\ \text{"strutture booleane" per un approfondimento} \end{array} \right)$

## Teorema

$\forall a, b \in \mathbb{Z}$   
 $b \neq 0 \Rightarrow \exists! (q, r) \in \mathbb{Z} \times \mathbb{N} \quad (a = bq + r \wedge r < |b|)$

DIM Esistenza:  $r := a \bmod b$

$$b \mid a - r \quad \exists q \in \mathbb{Z} \quad ((bq = a - r) \Leftrightarrow a = bq + r)$$

Unicità: se  $q' \neq r'$  sono come sopra e inoltre  $q' \in \mathbb{Z}$ ,  $r' \in \mathbb{N}$

$$a = bq' + r' \quad \& \quad r' < |b|$$

$$r \equiv_b r' \quad r, r' \in \{0, 1, \dots, |b|-1\}$$

$$r \equiv_b r \equiv_b r' \quad \rightarrow \quad r \equiv_b r' \quad \rightarrow \quad r = r'$$

$$bq + r = a = bq' + r' \Rightarrow bq = bq' \quad b \neq 0 \Rightarrow q = q'.$$

$[a]_m$  classe di resto modulo m

Se  $m \neq 0$ ,  $\equiv_m$  è nucleo di equivalenza di:  $n \in \mathbb{Z} \mapsto n \bmod m \in \mathbb{N}$

rest ( $n, m$ )

# Criteri per trovare il M.C.D.

$$\forall a, b \in \mathbb{Z}$$

$a|b \Rightarrow a$  è un M.C.D. ( $a, b$ )

Supponiamo:  $a = bq + r$  in  $\mathbb{Z}$

$$a = bq + r \Rightarrow \text{Div}_{\mathbb{Z}}(a) \cap \text{Div}_{\mathbb{Z}}(b) \cap \text{Div}_{\mathbb{Z}}(r)$$

$\forall d \in \mathbb{Z}$  d è un M.C.D. ( $a, b$ )  $\Leftrightarrow d$  è M.C.D. ( $b, r$ )

$$\forall n \in \mathbb{Z}$$

$$n|a \wedge n|b \Rightarrow n|r = a+b(-q)$$

$$n|b \wedge n|r \Rightarrow n|a = bq+r$$

Se  $r \neq 0$

$$(a, b) \mapsto (b, r) \quad a = bq + r \quad b = rq_1 + r_1 \quad |b| > r > r_1 > r_2$$

Se  $r_1 \neq 0$

$$(b, r) \mapsto (r, r_1) \quad r = r_1 q_{t+1} + r_{t+2} \quad \leftarrow \text{M.C.D. } (a, b)$$

Prima o poi, il resto  $= 0$   $r_{t+1} = r_{t+2} q_{t+3}$

Esempio:

MCD (140, 94):

$$140 = 94 \cdot 1 + 46$$

$$94 = 46 \cdot 2 + 0$$

$$46 = 23 \cdot 2 + 0$$

$$23 = 11 \cdot 2 + 1$$

$$11 = 1 \cdot 11 + 0 \quad \text{MCD} = 1$$

$$\forall k, u, v \in \mathbb{Z}$$

$$k \neq 0 \Rightarrow (u|v \Leftrightarrow ku|(kv))$$

oppure

$$\frac{140}{2} = 85 \quad \wedge \quad \frac{94}{2} = 47$$

$$85 = 47 \cdot 1 + 38$$

$$47 = 38 \cdot 1 + 9$$

$$38 = 9 \cdot 4 + 2$$

$$9 = 2 \cdot 4 + 1 \quad \leftarrow \text{MCD} = 1 \cdot 2$$

oppure:

$$85 = 47 \cdot 2 + (-9)$$

$$47 = (-9)(-5) + 2$$

$$-9 = 2 \cdot (-5) + 1$$



# Teorema di Bézout

$\forall a, b \in \mathbb{Z}$  Detto  $d$  un MCD ( $a, b$ )

1)  $\exists u, v \in \mathbb{Z} \quad (d = au + bv)$

Nel caso precedente:

$$1 = (-9) + 2 \cdot 5 \quad 2 = 47 + (-9) \cdot 5 \quad -9 = 85 + 47 \cdot (-2)$$

$$1 = \underline{-9} + \underline{2} \cdot 5 = \underline{-9} + (\underline{47} \cdot \underline{-9}) \cdot 5 = \frac{\cancel{-9} \cdot 26}{1+(-5) \cdot 5} + \underline{47} \cdot 5 = (\cancel{85} + \cancel{47} \cdot (-2)) \cdot 26 + \cancel{47} \cdot 5 = \underline{85} \cdot 26 + \underline{47} \cdot (-47)$$

Ci sono infinite coppie  $(u, v)$  tali che  $d = au + bv$  con  $b \neq 0 \neq a$

## Lemma di Euclide

$\forall a, b, c \in \mathbb{Z}$

$$(a|bc \wedge a \nmid b \text{ sono coprimi}) \Rightarrow a|c$$

DIM  $\exists u, v \in \mathbb{Z}$

$$1 = au + bv \quad c = \underbrace{acu}_{a|} + \underbrace{bcv}_{a|} \Rightarrow a|c$$

2)  $d \mathbb{Z} = \{eu + bv \mid u, v \in \mathbb{Z}\}$

DIM  $\forall k \in \mathbb{Z}$  Se  $d = au + bv \quad dk = a(uk) + b(vk) \in$   
 $d|a \wedge d|b \Rightarrow \forall u, v \in \mathbb{Z} \quad d|au + bv \in$

3)  $\forall c \in \mathbb{Z}$  l'equazione di Bézout  $ax + by = c$   
ha soluzione se e solo se  $d|c$

DIM  $d = au + bv \quad c = dk = a(uk) + b(vk)$

4)  $a \wedge b$  sono coprimi se e solo se  $\exists u, v \in \mathbb{Z} \quad (1 = au + bv)$

## Equazione congruenziale

$\forall m \in \mathbb{N}^*$

$\forall A, C \in \mathbb{Z}_m \quad AX = C \quad : \star$

Siano  $a \in A$  e  $c \in C$

$$A = [a]_m \quad C = [c]_m$$

$$\text{Allora } \star \rightarrow [a]_m X = [c]_m$$

Se  $u \in \mathbb{Z}$   $[u]_m$  è solutore se e solo se  $[a]_m [u]_m = [c]_m$ , cioè  $[au]_m = [c]_m$ , cioè  $au \equiv_m c$ .

\*\*:  $aX \equiv_m c$  equazione congruenziale: le sue soluzioni sono  $u \in \mathbb{Z}$  tali che  $au \equiv_m c$ , cioè tali che  $[u]_m$  sia soluzione di  $AX = C$ .

L'insieme delle soluzioni è un'unione di classi di resto modulo  $m$ . (o si vuole o infinito)

- $\forall u \in \mathbb{Z}$   $u$  è soluzione di \*\* se e solo se  $\exists v \in \mathbb{Z} \quad (u, v)$  è soluz. dell'equazione  $ax+my=c$ )

DIM  $u$  soluzione di \*\*  $\Leftrightarrow m | c - au \Leftrightarrow \exists v \in \mathbb{Z} \quad (mv = c - au)$

$\Leftrightarrow \exists v \in \mathbb{Z} \quad (au + mv = c) \Leftrightarrow \exists v \in \mathbb{Z} \quad (u, v)$  è soluzione di  $ax + by = c$ )

# Algoritmo euclideo, massimo comun divisore ed equazioni diofantee

Se  $a$  e  $b$  sono numeri interi, si dice che  $a$  divide  $b$ , in simboli:  $a \mid b$ , se e solo se esiste  $c \in \mathbb{Z}$  tale che  $b = ac$ . Si può subito notare che:

- 1 e  $-1$  sono gli unici interi che dividano ogni intero;
- 0 è l'unico intero che sia diviso da ogni intero.
- $\forall a, b \in \mathbb{Z} (a \mid b \iff -a \mid b \iff a \mid -b \iff -a \mid -b)$

L'insieme dei divisori (in  $\mathbb{Z}$ ) di un intero  $n$  si indica come  $D(n)$ . Dunque, per ogni  $n \in \mathbb{Z}$ ,

$$D(n) := \{a \in \mathbb{Z} : a \mid n\},$$

ad esempio,  $D(6) = \{1, -1, 2, -2, 3, -3, 6, -6\}$ .

Un *massimo comun divisore* tra  $a$  e  $b$  è poi un intero  $d$  per il quale valgono le due condizioni:

- i.)  $d \mid a \wedge d \mid b$ ; e
- ii.)  $\forall c \in \mathbb{Z} ((c \mid a \wedge c \mid b) \Rightarrow c \mid d)$ ;

ovvero, in modo equivalente:

- i.)  $d \in D(a) \cap D(b)$ ; e
- ii.)  $\forall c \in D(a) \cap D(b), c \mid d$ .

Dunque, un massimo comun divisore tra  $a$  e  $b$  è un divisore comune ad  $a$  e  $b$  che sia diviso da ogni altro divisore comune ad  $a$  e  $b$ .

Alcune osservazioni immediate sulla nozione di massimo comun divisore sono le seguenti:

- Se  $d$  è un massimo comun divisore tra  $a$  e  $b$  allora  $d$  e  $-d$  sono gli unici massimi comuni divisori tra  $a$  e  $b$ ,

dunque: calcolare un massimo comun divisore tra due interi equivale a calcolarli tutti;

- per ogni  $a, b \in \mathbb{Z}$ , i divisori comuni ad  $a$  e  $b$  sono tutti e soli i divisori comuni ad  $|a|$  e  $|b|$ ; quindi i massimi comuni divisori tra  $a$  e  $b$  sono tutti e soli i massimi comuni divisori tra  $|a|$  e  $|b|$ .

Quest'ultima osservazione mostra che nel calcolare massimi comuni divisori tra numeri interi è sempre possibile ridursi a calcolare massimi comuni divisori tra interi non negativi. Ad esempio, i massimi comuni divisori tra  $-7811$  e  $8456985$  sono precisamente i massimi comuni divisori tra  $7811$  e  $8456985$ , così come quelli tra  $-7811$  e  $-8456985$  o quelli tra  $7811$  e  $-8456985$ . Inoltre, il calcolo dei massimi comuni divisori tra  $0$  ed un arbitrario intero è immediato, come segue da queste altre due osservazioni:

- se  $a$  e  $b$  sono interi e  $a \mid b$ , allora  $a$  è un massimo comun divisore tra  $a$  e  $b$ ;
- in particolare, per ogni  $a \in \mathbb{Z}$ , si ha che  $a$  è un massimo comun divisore tra  $a$  e  $0$ .

Pertanto:

*Il problema di calcolare un massimo comun divisore tra numeri interi si riduce sempre al problema di calcolare un massimo comun divisore tra numeri interi positivi.*

Sino a questo momento non si è ancora stabilito se questo problema abbia sempre soluzione, cioè se, assegnati comunque due interi  $a$  e  $b$  esista un massimo comun divisore tra  $a$  e  $b$ .

Il teorema fondamentale dell'aritmetica suggerisce un metodo per calcolare un massimo comun divisore tra  $a$  e  $b$ , quello che viene insegnato sin dalla scuola elementare: supponendo, come lecito,  $a$  e  $b$  positivi, basta esprimere sia  $a$  che  $b$  come prodotti di potenze di numeri primi (positivi) a due a due distinti, con esponenti positivi:

$$\begin{aligned} a &= p_1^{\lambda_1} p_2^{\lambda_2} \cdots p_s^{\lambda_s} \\ b &= q_1^{\mu_1} q_2^{\mu_2} \cdots q_t^{\mu_t}; \end{aligned}$$

si ottiene un massimo comun divisore tra  $a$  e  $b$  come prodotto di tutti i primi che appaiono in entrambe le fattorizzazioni (i *fattori comuni* ...), ciascuno elevato al minimo degli esponenti con cui

---

**Avvertenza:** Queste note integrano, ma non sostituiscono, le corrispondenti parti del libro di testo.

appare (... col minimo esponente). Ciò è facile da verificare (lo si faccia per esercizio) e mostra che la risposta alla domanda formulata sopra è positiva. Grazie anche alle osservazioni precedenti possiamo concludere che:

*Se  $a$  e  $b$  sono interi allora:*

- se  $a = b = 0$ , l'unico massimo comun divisore tra  $a$  e  $b$  è 0;
- altrimenti, se almeno uno tra  $a$  e  $b$  è diverso da zero,  $a$  e  $b$  hanno esattamente due massimi comun divisori, uno opposto dell'altro.

*In ogni caso, dunque, esiste uno ed un solo massimo comun divisore non negativo tra  $a$  e  $b$ .*

Il massimo comun divisore non negativo tra due interi  $a$  e  $b$  viene spesso indicato con il simbolo  $\text{MCD}(a, b)$ .

Il metodo di calcolo di un massimo comun divisore tra due interi  $a$  e  $b$  appena ricordato è molto rapido ed efficace nel caso in cui  $a$  e  $b$  siano numeri di valore assoluto sufficientemente piccolo da renderne semplice la scomposizione in fattori primi. Quando si ha a che fare con numeri più grandi questo metodo risulta invece spesso impraticabile, dal momento che non sono noti metodi che permettono di scomporre in tempi ragionevolmente brevi numeri interi arbitrari; anzi, il calcolo dei fattori primi di un intero può rivelarsi di estrema complessità computazionale.

Per questo è molto importante disporre di un metodo alternativo, quello fornito dall'algoritmo euclideo, che ora illustreremo e che si dimostra essere invece molto efficiente. Per semplificare la discussione, introduciamo una definizione. Per ogni  $a, b \in \mathbb{Z}$  chiamiamo *combinazione lineare di  $a$  e  $b$  a coefficienti in  $\mathbb{Z}$*  ogni numero intero che si possa scrivere come  $\alpha a + \beta b$  per opportuni  $\alpha, \beta \in \mathbb{Z}$ . In altri termini, una combinazione lineare di  $a$  e  $b$  (a coefficienti in  $\mathbb{Z}$ ; talvolta lasceremo sottintesa questa specificazione) è la somma di un multiplo di  $a$  per un multiplo di  $b$ . Ad esempio, sono combinazioni lineare di  $a$  e  $b$  i numeri  $3a + 7b$ ,  $15a - 2b$ ,  $-19b$ .

**Lemma 1.** *Siano  $a, b, c \in \mathbb{Z}$ . Se  $c$  divide  $a$  e  $b$  allora  $c$  divide ogni combinazione lineare di  $a$  e  $b$  a coefficienti in  $\mathbb{Z}$ .*

*Dimostrazione* — Se  $c | a$  e  $c | b$ , esistono interi  $h$  e  $k$  tali che  $a = hc$  e  $b = kc$ . Scelti comunque  $\alpha, \beta \in \mathbb{Z}$  si ha allora  $\alpha a + \beta b = \alpha(hc) + \beta(kc) = (\alpha h + \beta k)c$ , dunque  $c$  divide  $\alpha a + \beta b$ .  $\square$

Un caso particolare del precedente lemma è il punto centrale del ragionamento che suggerisce e giustifica l'algoritmo euclideo:

**Lemma 2.** *Siano  $a, b, q, r \in \mathbb{Z}$  tali che  $a = bq + r$ . Allora i divisori comuni ad  $a$  e  $b$  sono tutti e soli i divisori comuni a  $b$  e  $r$ . In particolare, i massimi comun divisori tra  $a$  e  $b$  sono precisamente i massimi comun divisori tra  $b$  e  $r$ .*

*Dimostrazione* — Sia  $c$  un divisore comune a  $b$  e  $r$ . Poiché  $a$  è combinazione lineare di  $b$  e  $r$ , allora  $c | a$  per il Lemma 1. Dunque  $c$  è un divisore comune ad  $a$  e  $b$ . Abbiamo così provato l'inclusione

$$D(a) \cap D(b) \supseteq D(b) \cap D(r).$$

Per provare l'inclusione opposta, osserviamo che  $r = a - bq$  è combinazione lineare di  $a$  e  $b$ , quindi, come per il passaggio precedente, ogni divisore comune ad  $a$  e  $b$  divide  $r$  ed è così un divisore comune a  $b$  e  $r$ . Abbiamo ora dimostrato l'uguaglianza  $D(a) \cap D(b) = D(b) \cap D(r)$ , cioè che  $a$  e  $b$  da una parte e  $b$  ed  $r$  dall'altra hanno gli stessi divisori comuni, quindi anche gli stessi massimi comun divisori.  $\square$

Supponiamo ora di voler calcolare un massimo comun divisore tra due interi  $a$  e  $b$ ; come visto sopra possiamo supporre che essi siano entrambi positivi. Possiamo ovviamente anche supporre  $a \geq b$ , infatti se  $a < b$  basta scambiare tra loro  $a$  e  $b$ , dal momento che  $\text{MCD}(a, b) = \text{MCD}(b, a)$ .

Come sappiamo, si può effettuare la divisione aritmetica (con resto) di  $a$  per  $b$ . Esistono dunque (e sono univocamente determinati) due numeri naturali  $q$  (il quoziente) e  $r$  (il resto) tali che

$$a = bq + r \quad \text{e} \quad r < b.$$

Il Lemma 2 mostra che vale l'uguaglianza  $\text{MCD}(a, b) = \text{MCD}(b, r)$ . Possiamo dunque tradurre il nostro problema originale (calcolare un massimo comun divisore tra  $a$  e  $b$ ) con il problema, simile,

di calcolare un massimo comun divisore tra  $b$  e  $r$ . Il vantaggio di questa riformulazione consiste in questo, che se consideriamo la “grandezza” dei due numeri  $a$  e  $b$  come misura (grossolana!) della difficoltà del nostro problema (nel senso che è, probabilmente, più facile calcolare un massimo comun divisore tra due numeri più piccoli piuttosto che tra due numeri più grandi), allora l’aver sostituito la coppia  $(b, r)$  alla coppia  $(a, b)$  ha semplificato il problema, perché  $b < a$  e  $r < b$ .

È possibile che si abbia  $r = 0$ . In questo caso,  $b \mid a$  e quindi  $b$  è un massimo comun divisore tra  $a$  e  $b$ . Se invece  $r > 0$ , possiamo ripetere per  $b$  e  $r$  il procedimento effettuato per  $a$  e  $b$ : dividendo  $b$  per  $r$  otteniamo,

$$b = r_1 q_1 + r_1 \quad \text{e} \quad r_1 < r,$$

dove, ancora,  $q_1, r_1 \in \mathbb{N}$  e i massimi comuni divisori tra  $r$  e  $r_1$  sono i massimi comuni divisori tra  $b$  e  $r$ , quindi tra  $a$  e  $b$ . Se  $r_1 = 0$  (cioè se  $r \mid b$ ), allora  $r$  è un massimo comun divisore tra  $a$  e  $b$ , in caso contrario possiamo effettuare un’altra divisione, quella tra  $r$  e  $r_1$ , ottenendo  $q_2, r_2 \in \mathbb{N}$  tali che:

$$r = r_1 q_2 + r_2 \quad \text{e} \quad r_2 < r_1,$$

se  $r_2 = 0$  allora  $r_1$  è il massimo comun divisore cercato, altrimenti si proseguirà dividendo  $r_1$  per  $r_2$ .

Dovrebbe essere a questo punto chiaro il procedimento: ad ogni passo si verifica se il resto  $r_t$  dell’ultima divisione effettuata:  $r_{t-2} = r_{t-1} q_t + r_t$ , è 0; in questo caso il penultimo resto  $r_{t-1}$  (vale a dire, l’ultimo resto diverso da 0, o, ancora, l’ultimo divisore) è il massimo comun divisore positivo tra  $a$  e  $b$ , se invece  $r_t \neq 0$  si effettua un’altra divisione, tra il divisore  $r_{t-1}$  ed il resto  $r_t$  della divisione precedente.

È ancora da chiarire un solo punto, cioè se questo procedimento termina, ovvero se, iterando questo procedimento, si perviene ad una divisione con resto 0. La risposta è affermativa. Infatti, la sequenza dei resti ottenuti nelle successive divisioni è strettamente decrescente:

$$b > r > r_1 > r_2 > r_3 > \dots \geq 0$$

e una sequenza strettamente decrescente di numeri naturali minori di  $b$  può avere al più  $b$  termini, dal momento che l’insieme  $\{n \in \mathbb{N} \mid b \geq n\}$  ha  $b$  elementi. Dunque  $r_t = 0$  per qualche  $t < b$ . Pertanto l’algoritmo termina, fornendo un massimo comun divisore tra  $a$  e  $b$ , dopo al più  $b$  divisioni (ad essere pedanti, si dovrebbe specificare che, affinché tutto ciò che è stato appena scritto abbia senso in ogni caso, si devono sottintendere le posizioni  $r_0 := r$  e  $r_{-1} := b$ ).

Notiamo che l’algoritmo euclideo appena descritto fornisce un’altra dimostrazione, costruttiva, dell’esistenza di un massimo comun divisore tra due arbitrari interi.

Possiamo riassumere la discussione precedente e schematizzare l’algoritmo euclideo per la ricerca di un massimo comun divisore come segue:

Assegnati due numeri interi  $a$  e  $b$ , si intende calcolare un massimo comun divisore  $d$  tra  $a$  e  $b$ .

- ① se uno tra  $a$  e  $b$  è 0, si pone  $d$  uguale all’altro e l’algoritmo termina. Altrimenti:
- ② si sostituiscono  $a$  e  $b$  con  $|a|$  e  $|b|$ , nell’ordine;
- ③ se  $a < b$  si scambiano tra loro  $a$  e  $b$ ;
- ④ a partire dalla divisione di  $a$  per  $b$  si effettuano divisioni aritmetiche successive, come sopra specificato, finché non si ottenga 0 come resto:

$$\begin{aligned} a &= b q &+& r \\ b &= r_1 q_1 &+& r_1 \\ r &= r_1 q_2 &+& r_2 \\ r_1 &= r_2 q_3 &+& r_3 \\ &\vdots \\ r_{t-3} &= r_{t-2} q_{t-1} &+& r_{t-1} \\ r_{t-2} &= r_{t-1} q_t &+& r_t \\ r_{t-1} &= r_t q_{t+1} &+& 0 \end{aligned}$$

dove  $b > r > r_1 > r_2 > \dots > r_t > 0$ . A questo punto, si pone  $d = r_t$  e l’algoritmo termina.

Anche se non essenziali per la comprensione dell'algoritmo, si possono fare alcune osservazioni marginali. Innanzitutto, il passo ② non è necessario in senso stretto, dal momento che è possibile effettuare divisioni aritmetiche anche tra interi negativi, o tra un positivo e un negativo. Tuttavia, è consigliabile eseguirlo per evitare inutili complicazioni di calcolo. Anche il passo ③ si sarebbe potuto omettere dalla descrizione dell'algoritmo, per un motivo di tipo diverso. Infatti, se  $a$  e  $b$  sono positivi e  $a < b$ , allora la divisione aritmetica di  $a$  per  $b$  fornisce quoziente 0 e resto  $a$ . Dunque se eseguiamo il passo ④ dell'algoritmo senza aver prima scambiato tra loro  $a$  e  $b$ , la prima divisione (di  $a$  per  $b$ ) fornisce  $r = a$  e quindi la seconda, quella tra  $b$  e  $r = a$ , è precisamente quella da cui saremmo partiti se avessimo eseguito il passo ③. Ciò mostra che l'unico scopo del passo ③ è quello di evitare una divisione inutile (ancorché banale).

Osservazione più rilevante, a proposito del passo ④, è che, come abbiamo già detto, l'algoritmo terminerà dopo al più  $|b|$  divisioni.

Come esempio di applicazione dell'algoritmo euclideo, supponiamo di voler calcolare il massimo comun divisore positivo  $d$  tra 2547 e -7431. Passando ai valori assoluti dei due numeri considerati, e tenendo conto che  $2547 < 7431$ , procediamo come al passo ④ dopo aver posto  $a = 7431$  e  $b = 2547$ . Eseguiamo dunque le divisioni:

$$\begin{array}{r}
 7431 = 2547[2] + 2337 \\
 \downarrow \quad \downarrow \\
 2547 = 2337[1] + 210 \\
 \downarrow \quad \downarrow \\
 2337 = 210[11] + 27 \\
 \downarrow \quad \downarrow \\
 210 = 27[7] + 21 \\
 \downarrow \quad \downarrow \\
 27 = 21[1] + 6 \\
 \downarrow \quad \downarrow \\
 21 = 6[3] + 3 \\
 \downarrow \quad \downarrow \\
 6 = 3[2]
 \end{array}$$

Come evidenziato dalle frecce in colore, ogni divisione successiva alla prima ha per dividendo e per divisore il divisore ed il resto della divisione precedente. Come si può capire, è importante, eseguendo l'algoritmo, non confondere i ruoli tra i successivi divisori (indicati sopra come  $b, r, r_1, \dots$ ) ed i successivi quozienti (indicati come  $q, q_1, q_2, \dots$ ). I primi vanno riutilizzati nella divisione seguente, i secondi no. Allo scopo di evitare questa possibile confusione può essere utile adottare qualche artificio grafico. In questo caso, i quozienti sono stati scritti tra parentesi quadre.

Tornando al nostro specifico esempio, poiché l'ultimo resto non nullo è 3, concludiamo che 3 è un massimo comun divisore tra 2562 e -7491.

Un'ulteriore osservazione su questo algoritmo è che esso può essere reso ancora più efficace da una piccola modifica. Infatti, l'algoritmo si basa su una ripetuta applicazione del Lemma 2, e nell'enunciato del Lemma 2 non è richiesto che gli interi  $q$  ed  $r$  siano proprio il quoziente ed il resto della divisione aritmetica di  $a$  per  $b$ . Ora, è possibile effettuare un altro tipo di divisione tra interi che rispetti la condizione " $a = bq + r$ " dell'ipotesi del Lemma 2:

**Lemma 3** (Divisione euclidea). *Siano  $a, b \in \mathbb{Z}$ . Se  $b \neq 0$  esistono  $q, r \in \mathbb{Z}$  tali che  $a = bq + r$  e  $|r| \leq |b|/2$ .*

*Dimostrazione* — Assegnati  $a$  e  $b$  come richiesto dall'enunciato, effettuiamo la divisione aritmetica tra  $a$  e  $b$ . Otteniamo così  $\bar{q}, \bar{r} \in \mathbb{Z}$  tali che  $a = b\bar{q} + \bar{r}$  e  $0 \leq \bar{r} < |b|$ . Se  $\bar{r} \leq |b|/2$ , allora abbiamo concluso la ricerca di  $q$  e  $r$ : basterà porre  $r = \bar{r}$  e  $q = \bar{q}$ , a questo punto avremo  $a = bq + r$  e  $|r| = r \leq |b|/2$ , come richieso dall'enunciato.

Se invece  $\bar{r} > |b|/2$ , allora si ha  $|b| - \bar{r} < |b| - (|b|/2) = |b|/2$ . In questo caso, poniamo  $r := \bar{r} - |b|$ . Poiché  $\bar{r} < |b|$  si ha allora  $r < 0$ , e quindi  $|r| = -r = |b| - \bar{r} < |b|/2$ . Dunque la condizione  $|r| \leq |b|/2$  è soddisfatta. Inoltre  $\bar{r} = |b| + r$ , dunque

$$a = b\bar{q} + \bar{r} = b\bar{q} + (|b| + r) = bq + r,$$

avendo posto  $q = \bar{q} + 1$  se  $b > 0$  (e quindi se  $|b| = b$ ) e  $q = \bar{q} - 1$  se  $b < 0$ . Con questa scelta di  $q$  e  $r$  le condizioni richieste dall'enunciato sono soddisfatte, e così il lemma è dimostrato.  $\square$

Ad esempio la divisione aritmetica di 14 per 5 dà quoziante 2 e resto 4; la divisione euclidea appena introdotta dà quoziante 3 (aumentato di 1 rispetto al precedente, perché il divisore 5 è positivo) e resto  $-1$ : infatti  $14 = 5[3] + (-1)$ .

Possiamo eseguire l'algoritmo euclideo per la ricerca di un massimo comun divisore effettuando quest'ultimo tipo di divisioni anziché quelle aritmetiche. Uno svantaggio (se così si può dire, anche questo sarebbe evitabile) è che eseguiremo divisioni anche tra numeri negativi, un significativo vantaggio è che la successione dei resti  $r, r_1, r_2, \dots$  verificherà le condizioni:

$$|r| \leq |b|/2; \quad |r_1| \leq |r|/2 \leq |b|/4; \quad |r_2| \leq |r_1|/2 \leq |b|/8; \dots,$$

che, per dirla in termini informali, garantiscono che la successione dei resti decrescerà, nella maggior parte dei casi, più rapidamente di quanto non accadeva con la versione originaria dell'algoritmo. Ciò significa che possiamo aspettarci di dover effettuare meno divisioni, e quindi di terminare più rapidamente l'algoritmo.

A titolo di esempio, si possono confrontare il procedimento seguito prima per il calcolo del massimo comun divisore tra 7431 e 2547 con una versione dello stesso calcolo eseguito effettuando divisioni eucleedee anziché aritmetiche:

$$\begin{array}{ll} 7431 = 2547[2] + 2337 & 7431 = 2547[3] + (-210) \\ 2547 = 2337[1] + 210 & 2547 = (-210)[-12] + 27 \\ 2337 = 210[11] + 27 & -210 = 27[-8] + 6 \\ 210 = 27[7] + 21 & 27 = 6[4] + 3 \\ 27 = 21[1] + 6 & 6 = 3[2] \\ 21 = 6[3] + 3 & \\ 6 = 3[2] & \end{array}$$

**Esercizio.** Si sarebbero potute eseguire le divisioni nella colonna di sinistra, senza alterare il risultato finale, tralasciando i segni ‘meno’ dei divisori, e quindi assicurando che tutte le divisioni fossero tra numeri positivi. Ad esempio, dopo la prima divisione:  $7431 = 2547[3] + (-210)$ , la seconda avrebbe potuto essere  $2547 = 210[12] + 27$ . Basandosi sul Lemma 2 e su osservazioni precedenti, spiegare perché questa procedura è sempre lecita.

### Equazioni diofantee

Oltre al calcolo dei massimi comuni divisori, l'algoritmo euclideo permette di risolvere un altro importante problema. Un'equazione diofantea è un'equazione in cui appaiano solo indeterminate e numeri interi che si intenda risolvere in  $\mathbb{Z}$ , cioè per la quale siano ammesse come soluzioni solo numeri interi.

Ci occupiamo qui di un particolare tipo di equazione diofantea: quella cosiddetta lineare a due indeterminate, cioè una equazione diofantea della forma

$$ax + by = c, \tag{\ddagger}$$

dove  $a, b$  e  $c$  sono numeri interi. Risolvere l'equazione  $(\ddagger)$  significa dunque trovare le coppie di *intere*  $(u, v)$ , che rendano vera l'uguaglianza se sostituiti a  $x$  e  $y$ , cioè tali che  $au + bv = c$ . Osserviamo subito che è possibile che la  $(\ddagger)$  non ammetta soluzioni. Ad esempio, per  $a = b = 0$  e  $c = 1$  otteniamo l'equazione  $0x + 0y = 1$  che, ovviamente, non ammette soluzioni. Facendo uso della terminologia introdotta sopra, è chiaro che  $(\ddagger)$  ammette soluzioni (interne) se e solo se  $c$  è combinazione lineare di  $a$  e  $b$  a coefficienti in  $\mathbb{Z}$ . Ciò permette di dimostrare la prima importante osservazione su questo genere di equazioni.

**Lemma 4.** Siano  $a, b, c \in \mathbb{Z}$ , e sia  $d$  un massimo comun divisore tra  $a$  e  $b$ . Se  $d$  non divide  $c$ , allora l'equazione diofantea  $ax + by = c$  non ammette soluzioni.

*Dimostrazione* — Supponiamo che l'equazione abbia soluzioni. Allora esistono  $u, v \in \mathbb{Z}$  tali che  $au + bv = c$ , dunque  $c$  è combinazione lineare di  $a$  e  $b$  a coefficienti in  $\mathbb{Z}$ . Dal Lemma 1 segue allora che  $d$  divide  $c$ . Dunque, se supponiamo che  $d$  non divida  $c$  dobbiamo trarre la conclusione che la nostra equazione non ha soluzioni.  $\square$

Ad esempio, l'equazione diofantea  $2x + 6y = 3$  non ha soluzioni. Ovviamente ciò significa che l'equazione non ha soluzioni intere; essa ha ovviamente soluzioni razionali (cioè in  $\mathbb{Q}$ ), ad esempio  $(0, 1/2)$  o  $(1, 1/6)$ , ma nessuna di esse è data da due numeri interi.

Vedremo come l'algoritmo euclideo permette non solo di dimostrare che vale anche l'implicazione inversa di quella stabilità nel Lemma 4, cioè, nelle stesse notazioni, che se  $d$  divide  $c$ , allora l'equazione diofantea  $ax + by = c$  ammette soluzioni, ma anche di trovare queste soluzioni.

A questo scopo, iniziamo a considerare un caso banale, quello in cui almeno uno tra i coefficienti  $a$  e  $b$  è 0. Se  $a = 0$ , allora l'equazione  $(\ddagger)$  si riduce a  $by = c$ . Dire che questa ha una soluzione intera equivale a dire che  $b$  divide  $c$ . Inoltre,  $b$  è un massimo comun divisore tra  $a (= 0)$  e  $b$ , quindi è vero, in questo caso, che l'equazione data ammette soluzioni se (e solo se, in accordo col Lemma 4) un massimo comun divisore tra  $a$  e  $b$  divide  $c$ . Naturalmente, sempre in questo caso, è semplicissimo determinare le soluzioni, qualora ne esistano: se  $b \neq 0$  esse sono tutte (e sole) le coppie  $(n, c/b)$  al variare di  $n$  in  $\mathbb{Z}$ , mentre ogni coppia di numeri interi è soluzione se  $b = 0$ .

In modo analogo si ragiona se  $b = 0$ .

Supponiamo allora che sia  $a$  che  $b$  siano diversi da zero. Eseguiamo le divisioni successive previste dall'algoritmo euclideo:

$$\begin{aligned} a &= bq &+& r \\ b &= r_1 q_1 &+& r_1 \\ r &= r_1 q_2 &+& r_2 \\ r_1 &= r_2 q_3 &+& r_3 \\ &\vdots \\ r_{t-3} &= r_{t-2} q_{t-1} &+& r_{t-1} \\ r_{t-2} &= r_{t-1} q_t &+& r_t \\ r_{t-1} &= r_t q_{t+1} \end{aligned}$$

Allora  $r_t$  è uno dei due massimi comun divisori tra  $a$  e  $b$ ; poniamo  $d := r_t$ . Per risolvere l'equazione diofantea  $(\ddagger)$  proveremo prima a risolvere l'equazione diofantea

$$ax + by = d. \tag{\dagger}$$

Come già osservato, risolvere quest'ultima equivale ad esprimere  $d$  come combinazione lineare di  $a$  e  $b$  a coefficienti in  $\mathbb{Z}$ . La divisione in cui  $d = r_t$  appare come resto permette di esprimere  $d$  come combinazione lineare dei due resti precedenti,  $r_{t-1}$  e  $r_{t-2}$ , infatti da  $r_{t-2} = r_{t-1} q_t + d$  traiamo  $d = r_{t-2} + [-q_t]r_{t-1}$ . La divisione precedente,  $r_{t-3} = r_{t-2} q_t + r_{t-1}$ , fornisce poi  $r_{t-1}$  come combinazione lineare di  $r_{t-3}$  e  $r_{t-2}$ , dando  $r_{t-1} = r_{t-3} + [-q_{t-1}]r_{t-2}$ . Se sostituiamo questa espressione per  $r_{t-1}$  nella espressione trovata prima per  $d$  otteniamo  $d = r_{t-2} + [-q_t]r_{t-1} = r_{t-2} + [-q_t](r_{t-3} + [-q_{t-1}]r_{t-2}) = [-q_t]r_{t-3} + [1 + q_t q_{t-1}]r_{t-2}$ , ed esprimiamo così  $d$  come combinazione lineare di  $r_{t-2}$  e  $r_{t-3}$ , i due resti precedenti  $r_{t-1}$ . Questi passaggi dovrebbero essere sufficienti a comprendere l'intero procedimento. Per comodità di espressione poniamo  $r_0 := r$ ,  $r_{-1} := b$  e  $r_{-2} := a$ . Ad ogni passo  $d$  è espresso come combinazione lineare (a coefficienti in  $\mathbb{Z}$ ) di due "resti" con pedici consecutivi, diciamo  $r_i$  e  $r_{i+1}$ ; dalla divisione di  $r_{i-1}$  per  $r_i$  si ottiene  $r_{i+1} = r_{i-1} + [-q_{i+1}]r_i$ . Sostituendo nell'espressione di  $d$   $r_{i+1}$  con, appunto,  $r_{i-1} + [-q_{i+1}]r_i$  si può scrivere  $d$  come combinazione lineare di  $r_{i-1}$  e  $r_i$ , i due "resti" precedenti, nell'ordine,  $r_i$  e  $r_{i+1}$ . Questo passaggio può essere reiterato finché non si ottenga un'espressione di  $d$  come combinazione lineare di  $a = r_{-2}$  e  $b = r_{-1}$ , cioè due interi  $\alpha$  e  $\beta$  tali che  $\alpha a + \beta b = d$ . Allora  $\alpha$  e  $\beta$  forniscono una soluzione dell'equazione  $(\dagger)$ . Da questa si trae facilmente una soluzione per l'equazione  $(\ddagger)$ . Infatti, avendo assunto per ipotesi che  $d$  divida  $c$ , si ha  $c = dh$  per un opportuno  $h \in \mathbb{Z}$ . Allora, ponendo  $u := h\alpha$  e  $v := h\beta$ , si ha

$$au + bv = ah\alpha + bh\beta = h(\alpha a + \beta b) = hd = c,$$

quindi  $u$  e  $v$  forniscono una soluzione di  $(\ddagger)$ .

Abbiamo in questo modo provato che l'equazione diofantea  $(\ddagger)$  ammette soluzioni se  $d | c$ . Ricordandoci del Lemma 4 possiamo dunque concludere col seguente teorema:

**Teorema 5** (Teorema di Bézout). *Siano  $a, b \in \mathbb{Z}$ , e sia  $d = \text{MCD}(a, b)$ . Allora l'equazione diofantea  $ax + by = c$  ammette soluzioni (in  $\mathbb{Z}$ ) se e solo se  $d$  divide  $c$ .*

Il Teorema di Bézout è spesso citato, in modo equivalente, anche in questa forma:

**Teorema 5\*.** Siano  $a, b \in \mathbb{Z}$ , e sia  $d = \text{MCD}(a, b)$ . Allora l'insieme  $\{\alpha a + \beta b \mid \alpha, \beta \in \mathbb{Z}\}$  delle combinazioni lineari di  $a$  e  $b$  a coefficienti in  $\mathbb{Z}$  coincide con l'insieme  $d\mathbb{Z} = \{dk \mid k \in \mathbb{Z}\}$  dei multipli di  $d$  in  $\mathbb{Z}$ .

Come già per la ricerca di un massimo comun divisore, grazie all'algoritmo euclideo, non solo abbiamo stabilito esattamente quando un'equazione diofantea del tipo a noi considerato ammette soluzioni, ma abbiamo anche individuato un metodo (piuttosto efficace) per determinarne una nel caso esista.

Ad esempio, supponiamo di voler trovare soluzioni dell'equazione diofantea

$$74x + 22y = 10.$$

In questo caso è evidente che  $\text{MCD}(74, 22) = 2$ ; poiché 2 divide 10 siamo certi che l'equazione ammette soluzioni. Benché già conosciamo il massimo comun divisore 2, per trovare una soluzione mediante l'algoritmo euclideo bisogna eseguire le divisioni successive:

$$\begin{aligned} 74 &= 22[3] + 8 \\ 22 &= 8[2] + 6 \\ 8 &= 6[1] + 2 \\ 6 &= 2[3] \end{aligned}$$

sino ad ottenere 2 come ultimo resto non nullo. Da queste uguaglianze ricaviamo:

$$\begin{aligned} 8 &= 74 + 22[-3] \\ 6 &= 22 + 8[-2] \\ 2 &= 8 + 6[-1]. \end{aligned}$$

A questo punto possiamo esprimere 2 come combinazione lineare di 74 e 22, mediante successive sostituzioni:

$$\begin{aligned} 2 &= 8 + 6[-1] \\ &= 8 + (22 + 8[-2])[-1] && (\text{sostituendo } 6) \\ &= 8 + 22[-1] + 8[2] && (\text{eseguendo i calcoli ...}) \\ &= 8[3] + 22[-1] && (\dots \text{e raccogliendo i coefficienti di } 8 \text{ e } 22) \\ &= (74 + 22[-3])[3] + 22[-1] && (\text{sostituendo } 8) \\ &= 74[3] + 22[-9] + 22[-1] && (\text{eseguendo i calcoli ...}) \\ &= 74[3] + 22[-10] && (\dots \text{e raccogliendo i coefficienti di } 22 \text{ e } 74). \end{aligned}$$

Abbiamo così ottenuto l'espressione di 2 cercata. Questa mostra che la coppia  $(3, -10)$  è soluzione dell'equazione diofantea  $74x + 22y = 2$ . Moltiplicando per 5 (cioè per  $10/2$ ) si ottiene  $74[15] + 22[-50] = 10$ , dunque la coppia  $(15, -50)$  è soluzione della nostra equazione diofantea  $74x + 22y = 10$ .

Alcune annotazioni: come è chiaro, il procedimento effettuato si sarebbe potuto semplificare in almeno due modi:

- avremmo potuto effettuare divisioni euclidee anziché aritmetiche, risparmiando qualche passaggio. In questo caso specifico, dividendo 22 per 8 avremmo potuto scrivere  $22 = 8[3] + (-2)$  piuttosto che  $22 = 8[2] + 6$ , risparmiando sia una divisione che una sostituzione nella seconda parte dell'algoritmo. Per quest'ultima avremmo infatti ottenuto:  $-2 = 22 + 8[-3] = 22 + (74 + 22[-3])[-3] = 22[10] + 74[-3]$  e quindi  $10 = 22[-50] + 74[15]$ , moltiplicando per  $-5 = 10/(-2)$ .
- avendo osservato che 2 divide 74, 22 e 10, avremmo potuto semplificare l'equazione dividendo tutti i coefficienti per 2 e ottenendo  $37x + 11y = 5$ , un'equazione equivalente alla precedente. Avremmo poi proceduto con calcoli analoghi a quelli effettuati sopra, ma facilitati perché applicati a numeri già divisi per due.

Una cosa molto importante da chiarire è che la soluzione trovata non è l'unica. Infatti, come è immediato verificare, per ogni intero  $k$  si ha  $74(15 + 22k) + 22(-50 - 74k) = 10$ , il che fornisce infinite soluzioni alla nostra equazione. In effetti, in generale, *ogni equazione diofantea del tipo che stiamo considerando (cioè lineare a due indeterminate) ha infinite soluzioni se ne ha almeno una*. Detto in altri termini: ha nessuna o infinite soluzioni. Ciò si può far seguire dalla teoria delle equazioni congruenziali lineari ad una indeterminata, che viene trattata in altre note. Ci limitiamo qui a stabilire il nesso tra equazioni diofantee e equazioni congruenziali:

**Lemma 6.** Siano  $a, b, c, u \in \mathbb{Z}$ . Allora  $u$  è soluzione dell'equazione congruenziale  $ax \equiv c \pmod{b}$  se e solo se esiste  $v \in \mathbb{Z}$  tale che  $(u, v)$  sia soluzione dell'equazione diofantea  $ax + by = c$ .

*Dimostrazione* — Se  $u$  è soluzione di  $ax \equiv c \pmod{b}$ , allora  $b$  divide  $au - c$ , quindi  $au - c = bk$  per un opportuno  $k \in \mathbb{Z}$ . Ma allora  $au - bk = c$ , dunque, ponendo  $v = -k$ , si ha  $au + bv = c$ , il che significa che  $(u, v)$  è soluzione di  $ax + by = c$ . Viceversa, se esiste  $v \in \mathbb{Z}$  tale che  $(u, v)$  sia soluzione di  $ax + by = c$ , cioè tale che  $au + bv = c$ , allora  $b$  divide  $bv = au - c$ , dunque  $au \equiv c \pmod{b}$  e  $u$  è soluzione di  $ax \equiv c \pmod{b}$ .  $\square$

Possiamo concludere, grazie al Lemma 6, che il problema di risolvere l'equazione diofantea  $ax + by = c$  è equivalente al problema di risolvere l'equazione congruenziale  $ax \equiv c \pmod{b}$ . Infatti, ogni soluzione  $(u, v)$  della prima fornisce immediatamente la soluzione  $u$  della seconda; viceversa, se  $u$  è soluzione della seconda, allora non solo esiste  $v \in \mathbb{Z}$  tale che  $(u, v)$  sia soluzione della prima, ma tale  $v$  è facile da determinare: basta risolvere l'equazione (ad una sola indeterminata)  $au + bv = c$ .

A titolo di esempio, torniamo alla nostra equazione  $74x + 22y = 10$ . Come abbiamo visto, la coppia  $(15, -50)$  ne fornisce una soluzione. Allora  $15$  è soluzione dell'equazione congruenziale  $74x \equiv 10 \pmod{22}$ . Utilizzando questa informazione, dalla teoria delle equazioni congruenziali deduciamo che l'insieme di tutte le soluzioni di  $74x \equiv 10 \pmod{22}$  è  $[15]_{11} = [4]_{11} = 4 + 11\mathbb{Z} = \{4 + 11k \mid k \in \mathbb{Z}\}$ . Pertanto le soluzioni (interne) di  $74x + 22y = 10$  saranno tutte e sole le coppie  $(u, v)$  tali che  $u = 4 + 11k$  e  $v$  sia soluzione di  $74u + 22y = 10$ , vale a dire:  $v = (10 - 74u)/22$ . Svolgendo tutti i calcoli, si ottiene  $(10 - 74u)/22 = (5 - 37(4 + 11k))/11 = -13 - 37k$ , quindi l'insieme delle soluzioni della nostra equazione è

$$\{(4 + 11k, -13 - 37k) \mid k \in \mathbb{Z}\}.$$

Per  $k = 0$  troviamo così la soluzione  $(4, -13)$ , mentre la soluzione  $(15, -50)$  che avevamo calcolato sopra si ottiene per  $k = 1$ .

Va infine menzionata una importante applicazione del Teorema di Bézout. Due interi  $a$  e  $b$  si dicono *coprimi* se e solo se  $1 = \text{MCD}(a, b)$ . È evidente che questa condizione equivale a richiedere che non esista alcun numero primo che divida sia  $a$  che  $b$ . Il Teorema di Bézout ha questo caso particolare (anch'esso talvolta chiamato Teorema di Bézout, in effetti il Teorema 5 si può dedurre da questo enunciato):

**Corollario 7.** Siano  $a, b \in \mathbb{Z}$ . Allora  $a$  e  $b$  sono coprimi se e solo se esistono  $u, v \in \mathbb{Z}$  tali che  $au + bv = 1$ .

*Dimostrazione* — Per il Teorema 5, esistono  $u, v \in \mathbb{Z}$  tali che  $au + bv = 1$  se e solo se  $\text{MCD}(a, b)$  divide 1. Poiché i soli divisori di 1 sono 1 e  $-1$ , questa condizione equivale a  $\text{MCD}(a, b) = 1$ , cioè a richiedere che  $a$  e  $b$  siano coprimi.  $\square$

**Proposizione 8.** Siano  $a$  e  $b$  due interi coprimi. Per ogni  $c \in \mathbb{Z}$ , se  $a \mid bc$  allora  $a \mid c$ .

*Dimostrazione* — Per il Teorema di Bézout (o per il Corollario 7) si ha  $1 = au + bv$  per opportuni  $u, v \in \mathbb{Z}$ . Moltiplicando per  $c$  otteniamo  $c = acu + bcv$ . Dunque  $c$  è combinazione lineare di  $a$  e  $bc$ ; poiché  $a$  divide  $a$  e, per ipotesi,  $bc$  allora il Lemma 1 prova che  $a$  divide  $c$ .  $\square$

# Algebra

## Lezione 05/12



# Coprimi e Divisori dello zero

$\forall a, m \in \mathbb{Z}$

$$[\alpha]_m \in \mathcal{U}(\mathbb{Z}_m) \iff \begin{array}{l} \alpha x \equiv_m 1 \\ \text{ha soluzioni} \end{array} \iff \text{sono coprimi}$$

$$\exists B \in \mathbb{Z}_m \quad ([\alpha]_m B = [1]_m)$$

$$\exists b \in \mathbb{Z} \quad ([\alpha]_m [b]_m = [1]_m) \iff \exists b \in \mathbb{Z} \quad ([\alpha b]_m = [1]_m) \iff \exists b \in \mathbb{Z} \quad (\alpha b \equiv 1)$$

detto d un MCD( $a, m$ ) d/1

L'equazione congruenziale  
 $\alpha x \equiv_m 1$  ha soluzioni.

Esempio:

$$\mathbb{Z}_{10} \quad \forall a \in \mathbb{Z} \quad \bar{a} = [\alpha]_{10}$$

$$\mathbb{Z}_{10} = \{ \bar{0}, \bar{1}, \bar{2}, \bar{3}, \bar{4}, \bar{5}, \bar{6}, \bar{7}, \bar{8}, \bar{9} \}$$

$$\mathcal{U}(\mathbb{Z}_{10}) = \{ \bar{1}, \bar{3}, \bar{7}, \bar{9} \}$$

Gli altri sono tutti divisori dello zero.

DIM Sia  $m \neq 0$  ( $m > 0$ ). Sia  $a \in \mathbb{Z}$ ,  $a$  non coprimo con  $m$

$$\exists t \in \mathbb{Z} \quad (t \geq 1 \wedge t | a \wedge t | m) \quad 1 < l = \frac{m}{t} < m$$

$$[\alpha]_m \cdot [\ell]_m = [\alpha l]_m \quad t | a \Rightarrow t | \alpha l = m \quad [\alpha l]_m = [\bar{0}]_m \neq [\ell]_m$$

Enunciato:

$\forall m \in \mathbb{Z} \setminus \{0\}$  sono equivalenti:

- 1)  $m$  è primo
- 2)  $\mathbb{Z}_m$  è un campo
- 3)  $\mathbb{Z}_m$  è un dominio di integrità

DIM

$m$  primo

$$2) \mathbb{Z}_m \text{ campo} \iff [\alpha]_m \neq [\bar{0}]_m \Rightarrow [\alpha]_m \in \mathcal{U}(\mathbb{Z}_m)$$

$$3) m \in \mathbb{Z}$$

$$\Rightarrow \begin{array}{l} \exists r, k \in \mathbb{N} \\ |m| > 1 \end{array} \quad \begin{array}{l} \exists r, k \in \mathbb{N} \\ |m| = rk \quad \wedge \quad 1 < k < m \end{array}$$

$$[rk]_m = [lm]_m = [\bar{0}]_m, \text{ ma } [r]_m \neq [\bar{0}]_m \neq [k]_m$$

$m$  non primo

## Ancora Equazioni congruenziali

\*:  $ax \equiv_m c \quad (a, c, m \in \mathbb{Z})$

Detto  $d$  un MCD( $a, m$ ), \* ha soluzioni  $\iff d | c$ .

Sia  $S$  l'insieme delle soluzioni di \*.  $S = \{n \in \mathbb{Z} \mid an \equiv_m c\}$   $\leftarrow$  unione di classi modulo  $m$ .

$$\forall n \in \mathbb{Z} \quad \forall n \in S$$

$$n \equiv_m n \Rightarrow an \equiv_m an = c$$

Una prima soluzione costituisce nel trasformarla in un'equazione di fronte a:  $ax + my = c$

Sia  $R$  anello commutativo unitario

$$A, C \in R$$

l'equazione  $AX = C$  ha:

- esattamente una soluzione ( $A^{-1}C$ ) se  $A \in U(R)$ .
- al massimo una soluzione se  $A$  è cancellabile.  $\rightarrow 2x=4$  ha soluzione, ma  $2x=3$  nessuna.
- nessuna o più di una soluzione se  $A$  è un divisore dello zero.

$$1) AU = C$$

DIM Se  $U$  è una soluzione e  $B \in R - \{O_R\}$  sia tale che  $AB = O_R$ .

Allora  $U+B$  è una soluzione diversa da  $U$ .  $A(U+B) = AU+AB = C+O_R = C$

\*\* :  $[a]_m X = [c]_m$  equazione in  $\mathbb{Z}_m$

*S* è una classe modulo  $m$  (cioè \*\* ha esattamente una soluzione in  $\mathbb{Z}_m$ )

$$\Leftrightarrow [a]_m \in U(\mathbb{Z}_m) \iff a \text{ ed } m \text{ sono coprimi.}$$

In tal caso, l'equazione congruenziale si dice **ridotta**.

## Regole di semplificazione

①  $\forall a, c \in \mathbb{Z} \quad a \equiv_m a \quad \wedge \quad c \equiv_m c \Rightarrow *$  equivale a:  $a \times \equiv_m c$

Esempio:

$$\bullet 7214x \equiv_{10} \dots \quad \Leftrightarrow 4x \equiv_{10} \dots$$

$$\bullet 687x \equiv_{688} \dots \quad \Leftrightarrow -x \equiv_{688}$$

②  $\forall k \in \mathbb{Z} - \{0_R\} \Rightarrow *$  equivale a:  $a \times k \equiv_{mk} ck$

DIM  $\forall k \in \mathbb{Z} - \{0_R\} \quad \forall t, u \in \mathbb{Z} \quad t \equiv_m u \quad tk \equiv_{mk} uk$   
 $m | t-u \Leftrightarrow mk | (t-u)k$

③  $\forall t \in \mathbb{Z} \quad t \text{ è coprimo con } m \Rightarrow *$  equivale a  $a \times t \equiv_m ct$

DIM  $[a]_m [t]_m \times = [c]_m [t]_m$



I trovo un MCD ( $a, m$ ): d

II d divide c?  $\begin{cases} \text{no} & S=\emptyset \text{ STOP!} \\ \text{si} & \end{cases}$

III dividendo  $a, m, c$  per d: \* equivale a  $(\frac{a}{d})x \equiv_{\frac{m}{d}} c/d \leftarrow \text{ridotta!}$

IV  $\frac{\text{trovo}}{d} u \in \mathbb{Z}$  tale che  $[\frac{u}{d}]_m = [\frac{a}{d}]_m^{-1}$ , allora  $u \cdot (\frac{c}{d}) \in S \quad e \quad S = [\frac{uc}{d}]_m$   
 COME?

Ad esempio, così: algoritmo euclideo.

Risolvendo l'equazione congruenziale:  $(\frac{a}{d})x + (\frac{m}{d})y = 1$

se  $(\frac{a}{d})u + (\frac{m}{d})v = 1$ , allora  $u$  è l'intero cercato.

Esempi:

$$\bullet 2x \equiv_{10} 6 \quad S = [3]_5$$

$$\bullet 32x \equiv_8 8 \rightarrow 32x \equiv_8 12 \rightarrow 16x \equiv_8 6 \rightarrow 8x \equiv_8 3 \rightarrow -x \equiv_8 3 \quad S = [-3]_8$$

$$\bullet 14x \equiv_{11} 21 \rightarrow 2x \equiv_{11} 3 \rightarrow 2x \equiv_{11} 14 \rightarrow x \equiv_{11} 5 \quad S = [5]_{11}$$

$$\bullet 14x \equiv_{10} 21 \rightarrow \text{non ha soluzioni!} \quad 21 \text{ non è numero pari} \quad \text{e il modulo è pari.}$$

## Equazioni diofantee

$ax \equiv_m c \quad \forall u \in \mathbb{Z}, u \text{ è soluzione di } *$

se e solo se  $\exists v \in \mathbb{Z}$  tale che

$(u, v)$  è soluzione di  $ax + my = c$ .

L'insieme delle soluzioni delle diofantee è  $\{(u, v) \mid u \in S \wedge v = \frac{c-au}{m}\}$

$$au + mv = c \iff v = \frac{c-au}{m}$$



## Periodicità

$(G, \cdot)$  gruppo  $x \in G$

$\varphi: n \in \mathbb{Z} \mapsto x^n \in G$  omomorfismo da  $(\mathbb{Z}, +)$  a  $(G, \cdot)$

1)  $\varphi$  è iniettiva:  $\Leftrightarrow x$  ha periodo infinito / non è periodico

2)  $\varphi$  non è iniettiva:  $\Leftrightarrow x$  è periodico  $\Leftrightarrow A := \{n \in \mathbb{N}^* \mid x^n = 1_G\} \neq \emptyset$

$x$  periodico  $\Rightarrow \exists a, b \in \mathbb{Z} \quad (a > b \wedge x^{\frac{a}{a-b}} = x^b)$   
 $x^{\frac{a}{a-b}} = 1_G \quad a-b \in A$

Se  $x$  è periodico  $\circ(x) := \min A$  si chiama periodo di  $x$  in  $(G, \cdot)$ .

- $\circ(1_G) = 1$
- $\circ(-1)$  in  $(\mathbb{R} - \{0\}, \cdot)$  è 2
- 2 " " non è periodico
- $\forall x \quad \forall n \in \mathbb{N}^*$  in  $\text{Sym}(x)$  gli  $n$ -cili hanno periodo  $n$

Se  $n = \circ(x)$

$\forall k \in \mathbb{Z} \quad x^k = x^{k \mod n}$  infatti:  $\exists k \in \mathbb{Z} \quad r = k \mod n = nk + l$

$$x^k = x^{nk+l} = x^{nk} \cdot x^l = (x^n)^k \cdot x^l = 1_G^k \cdot x^l = x^l$$

$\forall a, b \in \mathbb{Z} \quad x^a = x^b \Leftrightarrow a \equiv_n b$

$\{x^t \mid t \in \mathbb{Z}\}$  ha esattamente  $n$  elementi.  
sottogruppo generato da  $x$

Algebra

Lezione 07/12



# POLINOMI SU UN ANELLO COMMUTATIVO UNITARIO

GIOVANNI CUTOLO

## 1. DEFINIZIONE E TERMINOLOGIA ESSENZIALE

Sia  $A$  un anello commutativo unitario. Per definizione, un *anello di polinomi* a coefficienti in  $A$  nell'indeterminata  $x$  è un anello commutativo unitario  $A[x]$  che verifichi le condizioni:

- (P<sub>1</sub>)  $A$  è un sottoanello unitario di  $A[x]$ ; <sup>(1)</sup>
- (P<sub>2</sub>)  $x \in A[x]$ ;
- (P<sub>3</sub>) per ogni  $f \in A[x]$  esiste una ed una sola successione  $\underline{a} = (a_i)_{i \in \mathbb{N}}$  di elementi di  $A$  con la proprietà che esista  $n \in \mathbb{N}$  tale che:
  - (i)  $f = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ , e
  - (ii)  $a_i = 0_A$  per ogni intero  $i > n$ .

È possibile dimostrare, ma non lo faremo qui, che per ogni anello commutativo unitario  $A$  esiste un anello di polinomi  $A[x]$  come qui specificato; vedremo [più avanti](#) che, fissato  $A$ , questi anelli di polinomi sono tutti isomorfi tra loro. Vengono chiamati *polinomi* (a coefficienti in  $A$ ) gli elementi di un tale anello  $A[x]$ .

Lasciando fisse le notazioni per  $A$  e  $x$  appena stabilite, facciamo alcune osservazioni che dovrebbero aiutare a comprendere meglio la definizione di anello di polinomi, introducendo nel contempo un po' di terminologia.

Come mostra (P<sub>1</sub>), gli elementi di  $A$  sono anch'essi polinomi; in questo contesto chiameremo spesso *polinomi costanti* gli elementi di  $A$ . Tra essi c'è  $0_A$  (che è ovviamente anche lo zero di  $A[x]$ ), chiamato anche *polinomio nullo*.

Per ogni  $f \in A[x]$ , la successione  $\underline{a} := (a_i)_{i \in \mathbb{N}}$  descritta in (P<sub>3</sub>) in relazione ad  $f$  viene chiamata la *successione dei coefficienti* di  $f$  e, per ciascun  $i \in \mathbb{N}$ , ci si riferisce talvolta ad  $a_i$  come al coefficiente di posto  $i$  (o coefficiente  $i$ -esimo) di  $f$ . Quanto richiesto al punto (ii) in (P<sub>3</sub>) mostra che l'insieme  $S_f := \{i \in \mathbb{N} \mid a_i \neq 0_A\}$  è finito (tutti gli elementi di  $S_f$  sono compresi tra 0 ed il numero che in (P<sub>3</sub>) appare come  $n$ ). Se  $f \neq 0_A$  si ha ovviamente  $S_f \neq \emptyset$ , quindi  $S_f$ , essendo un sottoinsieme finito non vuoto di  $\mathbb{N}$ , ha massimo; questo massimo è chiamato il *grado* di  $f$ , denotato col simbolo  $\nu f$  (o anche con altri simboli, tra i quali  $\nu(f)$ ,  $\deg f$  e  $\delta(f)$ ). Il coefficiente  $a_{\nu f}$  di posto  $\nu f$  si chiama *coefficiente direttore* di  $f$  e verrà indicato con  $\text{cd } f$ . In altri termini, se  $f$  è un polinomio non nullo, il grado di  $f$  è il massimo intero  $i$  tale che il coefficiente  $i$ -esimo di  $f$  non sia nullo, e questo stesso coefficiente è il coefficiente direttore di  $f$ . Ad esempio, in un anello  $\mathbb{Z}[x]$  di polinomi a coefficienti in  $\mathbb{Z}$  nell'indeterminata  $x$  c'è il polinomio  $h = 1 + 3x - 2x^3$ ; la successione dei coefficienti di  $h$  è la successione  $(a_i)_{i \in \mathbb{N}}$  di numeri interi definita ponendo  $a_0 = 1$ ,  $a_1 = 3$ ,  $a_3 = -2$  e  $a_i = 0$  per ogni  $i \in \mathbb{N} \setminus \{0, 1, 3\}$ . Allora  $S_h = \{0, 1, 3\}$ , quindi il grado di  $h$  è 3 ed il coefficiente direttore di  $h$  è  $a_3 = -2$ . Un altro esempio: se  $0_A \neq a \in A$  (cioè: se  $a$  è un polinomio costante non nullo) la successione  $(a_i)_{i \in \mathbb{N}}$  dei coefficienti di  $a$  è quella definita da  $a_0 = a$  ed  $a_i = 0_A$  per ogni  $i \in \mathbb{N}^*$ , dunque  $S_a = \{0\}$ , quindi  $a$  ha grado 0 e coefficiente direttore  $a$ .

Tornando al caso generale, le definizioni appena date di grado e di coefficiente direttore per un polinomio non nullo non si possono direttamente adattare al polinomio nullo  $0_A = 0_{A[x]}$  su un anello commutativo unitario  $A$ . Infatti, come si verifica immediatamente, il polinomio nullo ha la successione costante  $0_A$  (cioè la successione  $(a_i)_{i \in \mathbb{N}}$  in cui  $a_i = 0_A$  per ogni  $i \in \mathbb{N}$ ) come successione dei coefficienti. In altri termini: questo polinomio non ha coefficienti non nulli. Si conviene di estendere al polinomio nullo di  $A[x]$  le definizioni di coefficiente direttore e grado ponendo  $\text{cd } 0_A = 0_A$  e  $\nu 0_A = -\infty$ , dove  $-\infty$  è un simbolo, appunto, convenzionale al quale non attribuiamo uno specifico significato; richiediamo solo  $-\infty \notin \mathbb{N}$  ed estendiamo a  $\mathbb{N} \cup \{-\infty\}$  sia l'ordinamento che l'addizione usualmente definiti in  $\mathbb{N}$ , ponendo, per ogni  $n \in \mathbb{N} \cup \{-\infty\}$ ,  $-\infty \leq n$  e  $(-\infty) + n = n + (-\infty) = -\infty$ .

Osserviamo che, allora, i polinomi costanti sono tutti e soli i polinomi in  $A[x]$  di grado minore di 1: quelli non nulli hanno grado 0, mentre il polinomio nullo è l'unico che abbia grado  $-\infty$ ; tutti i polinomi non costanti hanno invece grado positivo. Inoltre, il polinomio nullo è l'unico polinomio con coefficiente direttore  $0_A$ .

Riassumendo: per ogni polinomio non nullo  $f \in A[x]$ , la proprietà (P<sub>3</sub>) garantisce che  $f$  si può scrivere in unico modo nella forma  $\sum_{i=0}^n a_i x^i$  dove  $n \in \mathbb{N}$ , per ogni  $i \in \{0, 1, 2, \dots, n\}$  si ha  $a_i \in A$  e  $a_n \neq 0_A$ ; questo accade se si pone  $n = \nu f$  e gli elementi  $a_i$  sono i primi  $n+1$  termini della successione dei coefficienti di  $f$  (e quindi  $a_n = \text{cd } f$ ); i termini rimanenti della successione dei coefficienti di  $f$  sono poi tutti uguali a  $0_A$ . <sup>(2)</sup>

Aggiungiamo ancora della terminologia: diremo che un polinomio è *monico* se e solo se il suo coefficiente direttore è  $1_A$ . Si usa infine chiamare *termine noto* di un polinomio il suo coefficiente di posto 0.

Completiamo questa sezione introduttiva notando esplicitamente una facile conseguenza dalla proprietà (P<sub>3</sub>). Se  $A$  è un anello commutativo unitario *non nullo* (cioè tale che  $A \neq \{0_A\}$ ), per ogni  $n \in \mathbb{N}$  il polinomio  $x^n$  ha

<sup>(1)</sup>si ricorda cosa questo significhi:  $A$  è un sottoanello di  $A[x]$  e l'unità  $1_A$  di  $A$  appartiene ad  $A[x]$ , quindi è anche l'unità di  $A[x]$ .

<sup>(2)</sup>possiamo anche aggiungere che, invece, verificano le proprietà richieste per  $n$  in (i) e (ii) di (P<sub>3</sub>) tutti e soli i numeri naturali  $n \geq \nu f$ .

grado  $n$  (perché il suo coefficiente  $n$ -esimo è  $1_A$  mentre tutti gli altri coefficienti sono uguali a  $0_A$ ), quindi  $x^n \notin A$  se  $n > 0$  (in particolare,  $x \notin A$ ) e se  $m$  è un numero naturale diverso da  $n$ , allora  $x^n \neq x^m$ . In altri termini: se  $|A| \neq 1$  le potenze di  $x$  in  $A[x]$  con esponente un numero naturale sono a due a due distinte, quindi: *se  $A$  è un anello commutativo unitario non nullo, l'anello di polinomi  $A[x]$  è infinito.*<sup>(3)</sup>

**Approfondimenti.**<sup>(4)</sup> Abbiamo un modo più preciso per chiarire la relazione che intercorre tra un polinomio e la sua successione dei coefficienti. Chiamiamo supporto di una successione  $\underline{a} := (a_i)_{i \in \mathbb{N}}$  di elementi di  $A$  l'insieme  $\{i \in \mathbb{N} \mid a_i \neq 0_A\}$ , e indichiamo con  $A_\omega$  l'insieme delle successioni di elementi di  $A$  con supporto finito. Le successioni dei coefficienti dei polinomi hanno supporto finito; abbiamo così l'applicazione  $\sigma: A[x] \rightarrow A_\omega$  che ad ogni polinomio in  $A[x]$  associa la sua successione dei coefficienti. Questa applicazione è biettiva. Infatti, sia  $\underline{a} = (a_i)_{i \in \mathbb{N}} \in A_\omega$ . Se  $\{i \in \mathbb{N} \mid a_i \neq 0_A\}$  non è vuoto sia  $n$  il suo massimo, altrimenti poniamo  $n = 0$ . In entrambi i casi, è chiaro che il polinomio  $f_{\underline{a}} := \sum_{i=0}^n a_i x^i$  ha  $\underline{a}$  come successione dei coefficienti. È altrettanto chiaro che, per ogni  $f \in A[x]$ , se  $\underline{a}$  è la successione dei coefficienti di  $f$ , allora  $f = f_{\underline{a}}$ . Dunque, l'applicazione  $\underline{a} \in A_\omega \mapsto f_{\underline{a}} \in A[x]$  è l'inversa di  $\sigma$ .

Menzioniamo il fatto che una delle possibili costruzioni di un anello dei polinomi su  $A$  si ottiene proprio definendo due operazioni (di addizione e moltiplicazione) nell'insieme  $A_\omega$  che rendano questo un anello commutativo unitario in cui si può immergere  $A$ , in modo che  $A_\omega$  risulti un anello di polinomi ad una indeterminata a coefficienti in  $A$ .

## 2. PROPRIETÀ UNIVERSALE

La proprietà più importante degli anelli di polinomi è la seguente:

**Proprietà universale per anelli di polinomi ad una indeterminata.** *Sia  $A[x]$  un anello di polinomi nell'indeterminata  $x$  sull'anello commutativo unitario  $A$ . Si fissino un anello commutativo unitario  $B$  ed un omomorfismo  $\theta: A \rightarrow B$  di anelli unitari<sup>(5)</sup> e  $b \in B$ . Allora esiste uno ed un solo omomorfismo  $\theta^*: A[x] \rightarrow B$  di anelli unitari tale che  $x^{\theta^*} = b$  e  $\theta$  sia la restrizione di  $\theta^*$  ad  $A$ .*<sup>(6)</sup>

In altre parole, assegnati omomorfismi (di anelli commutativi unitari) come nel diagramma a sinistra (l'omomorfismo  $A \hookrightarrow A[x]$  è l'immersione di  $A$  in  $A[x]$ ) ed un arbitrario  $b \in B$ , esiste uno ed un solo omomorfismo  $\theta^*$  che renda commutativo il diagramma a destra mandando  $x$  in  $b$ :

$$\begin{array}{ccc} A & \xrightarrow{\theta} & B \\ & \searrow & \nearrow \\ & A[x] & \end{array} \quad \begin{array}{ccc} A & \xrightarrow{\theta} & B \\ & \searrow & \nearrow \theta^* \\ & A[x] & \xrightarrow{x \mapsto b} \end{array}$$

Diamo solo un cenno alla dimostrazione, che si può completare per esercizio. Dalla definizione di omomorfismo di anelli segue subito che  $\theta^*$  non può essere altro che l'applicazione

$$\theta^*: \sum_{i=0}^n a_i x^i \in A[x] \longmapsto \sum_{i=0}^n a_i^{\theta} b^i \in B$$

(qui gli  $a_i$  rappresentano elementi di  $A$ ); è da osservare che ogni elemento di  $A[x]$  si scrive nella forma indicata, e segue dalla (P<sub>3</sub>) che l'applicazione  $\theta^*$  è ben definita; si può poi verificare che essa è effettivamente un omomorfismo di anelli unitari e che manda  $x$  in  $b$ . In questo modo la proprietà universale è provata.

Vediamo alcune importanti applicazioni della proprietà universale:

- *Unicità dell'anello dei polinomi, a meno di isomorfismi.* Supponiamo che  $A[x]$  e  $A[y]$  siano anelli di polinomi ad una indeterminata sullo stesso anello (commutativo unitario)  $A$ , con indeterminate, rispettivamente,  $x$  e  $y$ . Applichiamo la proprietà universale scegliendo come  $\theta$  l'immersione  $A \hookrightarrow A[y]$  e, come  $b$ , l'elemento  $y$ . Otteniamo così un (unico) omomorfismo  $\alpha: A[x] \rightarrow A[y]$  tale che  $x^\alpha = y$  e la restrizione di  $\alpha$  ad  $A$  sia l'immersione, cioè  $a^\alpha = a$  per ogni  $a \in A$ . Poiché anche  $A[y]$  è un anello dei polinomi, possiamo ripetere la stessa costruzione scambiando i ruoli di  $A[x]$  (ed  $x$ ) e  $A[y]$  (ed  $y$ ),

$$\begin{array}{ccc} A & \xleftarrow{\quad} & A[y] \\ & \searrow & \nearrow \alpha \\ & A[x] & \xrightarrow{x \mapsto y} \end{array} \quad \begin{array}{ccc} A & \xleftarrow{\quad} & A[x] \\ & \searrow & \nearrow \beta \\ & A[y] & \xrightarrow{y \mapsto x} \end{array}$$

ottenendo un omomorfismo  $\beta: A[y] \rightarrow A[x]$  tale che  $y^\beta = x$  e  $a^\beta = a$  per ogni  $a \in A$ . È facile verificare che  $\alpha$  e  $\beta$  sono l'uno l'inverso dell'altro. Infatti, per ogni elemento  $f = \sum_{i=0}^n a_i x^i$  di  $A[x]$  si ha  $f^{\alpha \beta} = (\sum_{i=0}^n a_i y^i)^\beta = \sum_{i=0}^n a_i x^i = f$  e, similmente,  $g^{\beta \alpha} = g$  per ogni  $g \in A[y]$ . Ciò prova che  $\alpha$  è un isomorfismo.

<sup>(3)</sup>segue invece facilmente da (P<sub>1</sub>) che se  $A$  è l'anello nullo, cioè se  $|A| = 1$ , allora  $A[x] = A$ .

<sup>(4)</sup>non richiesti ai fini dell'esame.

<sup>(5)</sup>si intende con questo che  $\theta$  è un omomorfismo di anelli che manda l'unità di  $A$  nell'unità di  $B$ .

<sup>(6)</sup>in queste note le immagini di elementi mediante applicazioni sono generalmente indicate con la notazione esponenziale, quindi, ad esempio,  $x^{\theta^*}$  è l'immagine di  $x$  mediante  $\theta^*$ .

Dunque, assegnati due anelli di polinomi ad una indeterminata su  $A$  esiste un isomorfismo tra questi due anelli di polinomi che manda l'indeterminata del primo nell'indeterminata del secondo e manda in se stesso ogni elemento di  $A$ . Con le notazioni appena usate, questo isomorfismo è l'applicazione

$$\alpha: \sum_{i=0}^n a_i x^i \in A[x] \mapsto \sum_{i=0}^n a_i y^i \in A[y],$$

osserviamo esplicitamente che essa manda ogni polinomio  $f$  di  $A[x]$  nel polinomio di  $A[y]$  che ha la stessa successione dei coefficienti di  $f$ .

Possiamo dunque dire, in modo un pò approssimativo ma efficace, che due anelli di polinomi sullo stesso anello commutativo unitario  $A$  possono solo differire per il nome dell'indeterminata; in questo senso, a meno di isomorfismi, ne esiste solo uno. Per questo motivo possiamo decidere di aver fissato, per ogni scelta di  $A$ , un anello dei polinomi  $A[x]$  ad una indeterminata su  $A$  e chiamare questo *l'anello dei polinomi ad una indeterminata su  $A$*  (con l'articolo determinativo).

- L'applicazione più frequente della proprietà universale si ha per il caso in cui  $B = A$  e  $\theta$  è l'applicazione identica di  $A$ . In questo caso la proprietà ci dice che per ogni  $c \in A$  esiste uno ed un solo omomorfismo di anelli unitari  $A[x] \rightarrow A$  che manda ogni elemento di  $A$  in sé e  $x$  in  $c$ :

$$\begin{array}{ccc} A & \xrightarrow{\text{id}_A} & A \\ & \searrow & \swarrow x \mapsto c \\ & A[x] & \end{array}$$

È facile descrivere esplicitamente questo omomorfismo. Per ogni  $f = \sum_{i=0}^n a_i x^i \in A[x]$  poniamo  $f(c) = \sum_{i=0}^n a_i c^i$ . L'omomorfismo di cui stiamo parlando è allora l'applicazione:

$$f \in A[x] \mapsto f(c) \in A,$$

che chiamiamo *omomorfismo di sostituzione*.

- Un'applicazione più specifica: per ogni intero positivo  $m$ , sia  $\varepsilon_m: n \in \mathbb{Z} \mapsto [n]_m \in \mathbb{Z}_m$ , la proiezione canonica  $\mathbb{Z} \twoheadrightarrow \mathbb{Z}_m$  (il simbolo di freccia a doppia punta ci ricorda il fatto che  $\varepsilon_m$  è un omomorfismo suriettivo). Componendo questa con l'immersione  $\iota_m: \mathbb{Z}_m \hookrightarrow \mathbb{Z}_m[x]$  otteniamo l'omorfismo di anelli unitari<sup>(7)</sup>  $\varepsilon_m \iota_m: n \in \mathbb{Z} \mapsto [n]_m \in \mathbb{Z}_m[x]$  (come di consueto, usiamo lo stesso simbolo,  $x$ , per l'indeterminata di diversi anelli di polinomi). La proprietà universale fornisce l'omomorfismo  $\bar{\varepsilon}_m$  qui descritto:

$$\begin{array}{ccc} \mathbb{Z} & \xrightarrow{\varepsilon_m} & \mathbb{Z}_m \hookrightarrow \mathbb{Z}_m[x] \\ & \searrow & \swarrow \bar{\varepsilon}_m \\ & \mathbb{Z}[x] & \end{array}$$

(è facile verificare che  $\bar{\varepsilon}_m$  è effettivamente suriettivo). Più esplicitamente, l'immagine mediante  $\bar{\varepsilon}_m$  di  $f = \sum_{i=0}^n a_i x^i \in \mathbb{Z}[x]$  è il polinomio  $f_m = \sum_{i=0}^n [a_i]_m x^i$ . Per indicare questo polinomio scriviamo spesso  $\sum_{i=0}^n \bar{a}_i x^i \in \mathbb{Z}_m[x]$ ; (è ovviamente essenziale indicare sempre il modulo  $m$ , per evitare ambiguità) o anche, ancora più semplicemente,  $\sum_{i=0}^n a_i x^i \in \mathbb{Z}_m[x]$ . Si dice che  $f_m$  è *il polinomio  $f$  riguardato come polinomio a coefficienti in  $\mathbb{Z}_m$*  (o anche ... modulo  $m$ ). Ad esempio, se  $f = 14x^3 - 3x + 1 \in \mathbb{Z}[x]$ , il polinomio  $f$  riguardato come polinomio a coefficienti in  $\mathbb{Z}_5$  è  $\bar{4}x^3 - \bar{3}x + \bar{1} \in \mathbb{Z}_5[x]$ , che possiamo anche scrivere come  $-x^3 + \bar{2}x + \bar{1} \in \mathbb{Z}_5[x]$ , o in infiniti altri modi.

Possiamo riferirci a quest'ultima costruzione per illustrare con qualche esempio le nozioni introdotte nella sezione precedente e vedere come grado e coefficiente direttore possono cambiare nel passaggio da un polinomio a coefficienti in  $\mathbb{Z}$  al corrispondente polinomio riguardato modulo un intero positivo. Il polinomio  $f = 15x^4 + 6x^2 + 2 \in \mathbb{Z}[x]$  ha grado 4 e coefficiente direttore 15. Invece  $f_5$ , cioè  $f$  riguardato come polinomio a coefficienti in  $\mathbb{Z}_5$  ha grado 2 (il suo quarto coefficiente è  $[15]_5 = [0]_5 = 0_{\mathbb{Z}_5}$  e coefficiente direttore  $[6]_5 = [1]_5 = 1_{\mathbb{Z}_5}$ , dunque  $f_5$  è monico; possiamo scrivere  $f_5 = x^2 + \bar{2} \in \mathbb{Z}_5[x]$ . Invece  $f_3$ , cioè  $f$  riguardato come polinomio a coefficienti in  $\mathbb{Z}_3$ , ha grado 0 e si ha  $f_3 = \text{cd } f_3 = [2]_3 = -1_{\mathbb{Z}_3}$ ; dunque  $f_3$  è un polinomio costante.

### 3. GRADO DI SOMME E PRODOTTI DI POLINOMI

Siano, ancora,  $A$  un anello commutativo unitario, e siano  $f, g \in A[x]$ , con successioni dei coefficienti, rispettivamente,  $(a_n)_{n \in \mathbb{N}}$  e  $(b_n)_{n \in \mathbb{N}}$ . Supponiamo anche  $f \neq 0_A \neq g$  e poniamo  $n = \nu f$ ,  $m = \nu g$ ; dunque  $f = \sum_{i=0}^n a_i x^i$  e  $g = \sum_{i=0}^m b_i x^i$ ; inoltre  $a_n = \text{cd } f \neq 0_A$  e  $b_m = \text{cd } g \neq 0_A$ . Allora, posto  $M = \max\{n, m\}$ ,

$$f + g = \sum_{i=0}^M (a_i + b_i) x^i; \quad f - g = \sum_{i=0}^M (a_i - b_i) x^i; \quad fg = \sum_{i=0}^{n+m} (a_0 b_0 + (a_0 b_1 + a_1 b_0)x + \dots + (a_n b_m) x^{n+m}). \quad (8)$$

<sup>(7)</sup>In conformità all'uso della notazione esponenziale per le immagini di elementi del dominio di un'applicazione, la composizione di applicazioni è indicata in queste note nell'ordine naturale, scrivendo a sinistra l'applicazione che agisce per prima, quindi  $fg$  piuttosto che  $g \circ f$  se  $f$  e  $g$  sono applicazioni componibili. Ad esempio,  $\varepsilon_m \iota_m = \iota_m \circ \varepsilon_m$

<sup>(8)</sup>più precisamente:  $fg = \sum_{i=0}^{n+m} c_i x^i$ , dove, per ogni  $i$ , si ha  $c_i = \sum_{j=0}^i a_j b_{i-j} = a_0 b_i + a_1 b_{i-1} + a_2 b_{i-2} + \dots + a_{i-1} b_1 + a_i b_0$ .

Cosa possiamo dire sui gradi di questi tre polinomi? Consideriamo in primo luogo  $f + g$ . Nella sua espressione non appaiono potenze di  $x$  con esponente superiore a  $M$ , quindi certamente  $\nu(f + g) \leq M$ , e  $\nu(f + g) = M$  se e solo se il coefficiente di posto  $M$  in  $f + g$  (cioè  $a_M + b_M$ ) è diverso da zero. Distinguiamo tre casi: se  $n < m$  allora  $M = m$  e  $a_m = 0_A$ , quindi  $a_M + b_M = b_m \neq 0_A$ . In questo caso, dunque,  $\nu(f + g) = m = M$ , inoltre  $\text{cd}(f + g) = b_m = \text{cd } g$ . Ad esempio: se  $A = \mathbb{Z}$ ,  $f = 2x + 1$  e  $g = 3x^4 + 2x + 2$  (quindi  $n = 1 < m = 4$ ) allora  $f + g = 3x^4 + 4x + 3$  ha grado  $4 = m$ . Similmente, se  $n > m$ , vediamo che  $f + g$  ha grado  $n$  e coefficiente direttore  $a_n = \text{cd } f$ . Nel terzo caso, quello in cui  $n = m$ , bisogna fare una distinzione ulteriore: se  $a_n + b_n \neq 0_A$  abbiamo  $\nu(f + g) = n = M$  e  $\text{cd}(f + g) = a_n + b_n$ , ma se invece  $a_n + b_n = 0_A$  (cioè  $a_n = -b_n$ ) allora certamente  $\nu(f + g) < n$ , perché per ogni intero  $i > n - 1$  il coefficiente  $i$ -esimo di  $f + g$  è  $0_A$ . Possiamo riassumere così ciò che abbiamo provato sino a questo punto:

**Proposizione 1.** Se  $A$  è un anello commutativo unitario e  $f, g \in A[x] \setminus \{0_A\}$ , allora  $\nu(f + g) = \max\{\nu f, \nu g\}$  a meno che  $\nu f = \nu g$  e  $\text{cd } f = -\text{cd } g$ . In questo secondo caso  $\nu(f + g) < \nu f = \nu g$ .

Ripetendo il ragionamento per  $f - g$ , oppure applicando la [Proposizione 1](#) a  $f$  e  $-g$  (perché  $f - g = f + (-g)$ ) si ha:

**Proposizione 2.** Se  $A$  è un anello commutativo unitario e  $f, g \in A[x] \setminus \{0_A\}$ , allora  $\nu(f - g) = \max\{\nu f, \nu g\}$  a meno che  $\nu f = \nu g$  e  $\text{cd } f = \text{cd } g$ . In questo secondo caso  $\nu(f - g) < \nu f = \nu g$ .

Vediamo qualche altro esempio, ancora in  $\mathbb{Z}[x]$ . Se  $f = 3x^2 + x + 1$  e  $g = 2x^2 + x + 2$  (quindi  $n = m = 2$ ) allora  $f - g = x^2 - 1$  ha anch'esso grado 2 ( $f$  e  $g$  hanno lo stesso grado, ma coefficienti direttori diversi); se invece  $g = 3x^2 + x + 2$  allora  $\nu f = \nu g$  e  $\text{cd } f = \text{cd } g$ , quindi  $f - g$  non ha grado 2, infatti  $f - g = -1$  ha grado 0.

Passiamo ora a considerare il grado di  $fg$ . Il ragionamento è simile: poiché nell'espressione di  $fg$  non appaiono potenze di  $x$  con esponente superiore a  $n + m$  certamente  $\nu(fg) \leq n + m$  e vale  $\nu(fg) = n + m$  se e solo se  $a_n b_m$  (il coefficiente  $(n + m)$ -esimo di  $fg$ ) è diverso da zero. Abbiamo allora:

**Proposizione 3.** Se  $A$  è un anello commutativo unitario e  $f, g \in A[x] \setminus \{0\}$ , posto  $a = \text{cd } f$  e  $b = \text{cd } g$  si ha:

- (i) se  $ab \neq 0_A$ , allora  $\text{cd}(fg) = ab$  e  $\nu(fg) = \nu f + \nu g$ . In particolare,  $fg \neq 0_A$ ;
- (ii) se  $ab = 0_A$ , allora  $\nu(fg) < \nu f + \nu g$ .

Se si verifica  $\nu(fg) = \nu f + \nu g$ , come nel caso (i) di questa proposizione, si dice che per i polinomi  $f$  e  $g$  vale la *regola di addizione dei gradi*. Ovviamente questa regola vale sempre nel caso in cui uno tra  $f$  e  $g$  è il polinomio nullo: se, ad esempio,  $f = 0_A$ , allora  $\nu(fg) = \nu(0_A) = -\infty = (-\infty) + \nu g = \nu f + \nu g$ .

Alcune importanti conseguenze della [Proposizione 3](#) sono:

**Corollario 4.** Sia  $A$  un anello commutativo unitario e sia  $f \in A[x]$ . Se  $\text{cd } f$  è cancellabile in  $A$  allora  $f$  è cancellabile in  $A[x]$  e, per ogni  $g \in A[x]$ , si ha  $\nu(fg) = \nu f + \nu g$ .

*Dimostrazione.* Sia  $g \in A[x] \setminus \{0_A\}$  e siano  $a = \text{cd } f$  e  $b = \text{cd } g$ . Poiché  $a$  è cancellabile in  $A$ , quindi non un divisore dello zero, e  $b \neq 0_A$  allora  $ab \neq 0_A$ . Per la [Proposizione 3](#), allora  $fg \neq 0_A$  e, inoltre,  $\nu(fg) = \nu f + \nu g$ . La prima informazione ci dice che  $f$  non è un divisore dello zero, e quindi è cancellabile, in  $A[x]$ . Con questo (ricordando che la regola di addizione dei gradi vale banalmente se uno dei polinomi coinvolti è quello nullo) la dimostrazione è completa.  $\square$

**Proposizione 5.** Sia  $A$  un anello commutativo unitario. Sono allora equivalenti:

- (i)  $A$  è un dominio di integrità;
- (ii) per ogni coppia di polinomi  $A[x]$  vale la regola di addizione dei gradi (cioè:  $\forall f, g \in A[x] (\nu(fg) = \nu f + \nu g)$ );
- (iii)  $A[x]$  è un dominio di integrità.

Inoltre, se  $A$  è un dominio di integrità allora  $\mathcal{U}(A[x]) = \mathcal{U}(A)$ .<sup>(9)</sup>

*Dimostrazione.* Supponiamo che  $A$  sia un dominio di integrità, e siano  $f, g \in A[x]$ . Se  $f = 0_A$ , allora vale banalmente  $\nu(fg) = \nu(0_A) = -\infty = \nu f + \nu g$ . Se invece  $f \neq 0_A$ , allora  $\text{cd } f \neq 0_A$ , quindi, poiché  $A$  è un dominio di integrità,  $\text{cd } f$  è cancellabile in  $A$  e  $\nu(fg) = \nu f + \nu g$  per il [Corollario 4](#). Abbiamo provato che (i) implica (ii).

Che (ii) implica (iii) è ovvio: se  $f, g \in A[x] \setminus \{0_A\}$  e in  $A[x]$  vale la regola di addizione dei gradi, allora, come per la [Proposizione 3](#), si ha  $\nu(fg) = \nu f + \nu g \geq 0$ , quindi  $fg \neq 0_A$ ; questo significa che in  $A[x]$  vale la legge di annullamento del prodotto e dunque  $A[x]$  è un dominio di integrità.

Anche l'implicazione (iii)  $\Rightarrow$  (i) è banale: se  $A[x]$  è un dominio di integrità e  $a$  e  $b$  sono elementi di  $A \setminus \{0_A\}$ , allora  $ab \neq 0_A$ , perché altrimenti  $a$  sarebbe un divisore dello zero in  $A[x]$ .<sup>(10)</sup>

Resta solo da provare che  $\mathcal{U}(A[x]) = \mathcal{U}(A)$  se valgono le condizioni (i), (ii) e (iii). Se  $a \in \mathcal{U}(A)$  e  $b$  è l'inverso di  $a$  in  $A$ , allora  $ab = 1_A = 1_{A[x]}$ , quindi  $b$  è anche l'inverso di  $a$  in  $A[x]$ , dunque  $a \in \mathcal{U}(A[x])$ . Pertanto  $\mathcal{U}(A) \subseteq \mathcal{U}(A[x])$ .<sup>(11)</sup> Nell'ipotesi che  $A$  sia un dominio di integrità sia, viceversa,  $f \in \mathcal{U}(A[x])$  e sia  $g$  l'inverso di  $f$  in  $A[x]$ . Allora  $fg = 1_A$  e, ovviamente,  $f \neq 0_A \neq g$ . Poiché in  $A[x]$  vale la regola di addizione dei gradi,

<sup>(9)</sup>ricordiamo che, per ogni anello unitario  $R$ , con  $\mathcal{U}(R)$  si indica il gruppo moltiplicativo degli elementi invertibili di  $R$ .

<sup>(10)</sup>in sostanza, ciò che stiamo osservando è che i sottoanelli non nulli dei domini di integrità sono essi stessi domini di integrità.

<sup>(11)</sup>anche in questo caso l'argomentazione mostra qualcosa che vale in contesti più generali: se  $A$  è un sottoanello *unitario* di un anello unitario  $R$ , allora  $\mathcal{U}(A) \subseteq \mathcal{U}(R)$ .

$\nu f + \nu g = \nu(fg) = \nu(1_A) = 0$ . Dunque,  $\nu f$  e  $\nu g$  sono due numeri naturali la cui somma è 0; di conseguenza  $\nu f = \nu g = 0$ . Ciò mostra che  $f \in A$  e  $g \in A$ , quindi sia  $f$  che il suo inverso sono elementi di  $A$ , dunque  $f \in \mathcal{U}(A)$ . Abbiamo così provato anche l'inclusione  $\mathcal{U}(A[x]) \subseteq \mathcal{U}(A)$ ; a questo punto la dimostrazione è completa.  $\square$

Vediamo così che la regola di addizione dei gradi non vale per polinomi su anelli che non siano domini di integrità, ed è importante osservare che per tali anelli può non valere neanche la conclusione finale della [Proposizione 5](#): non è detto che i polinomi invertibili siano costanti. Se ad esempio scegliamo come  $A$  l'anello  $\mathbb{Z}_4$  e poniamo  $f = \bar{2}x + \bar{1} \in \mathbb{Z}_4[x]$ , allora  $f^2 = \bar{4}x^2 + \bar{4}x + \bar{1} = \bar{1} = 1_{\mathbb{Z}_4}$ , quindi non solo  $0 = \nu(f \cdot f) < \nu f + \nu f$  e non vale in questo caso la regola di addizione dei gradi, ma addirittura abbiamo  $f \in \mathcal{U}(\mathbb{Z}_4[x])$  (si ha  $f^{-1} = f$ ) pur non essendo  $f$  un polinomio costante. Quindi  $\mathcal{U}(\mathbb{Z}_4[x]) \neq \mathcal{U}(\mathbb{Z}_4)$ , anzi  $\mathcal{U}(\mathbb{Z}_4[x]) \not\subseteq \mathbb{Z}_4$ .

Un ragionamento simile a quello svolto nella dimostrazione della [Proposizione 5](#) utilizzando la regola di addizione dei gradi mostra che i polinomi monici di grado maggiore di zero non sono mai invertibili. Ad esempio, qualunque sia l'anello commutativo unitario non nullo  $A$ , l'indeterminata  $x$  non è invertibile in  $A[x]$ . Infatti, se  $x$  fosse invertibile, detto  $g$  il suo inverso, avremmo  $1_A = xg$  e quindi  $\nu(xg) = 0$ , ma, per il [Corollario 4](#), vale per  $x$  e  $g$  la regola di addizione dei gradi, quindi  $\nu(xg) = \nu x + \nu g = 1 + \nu g$ , in contraddizione con quanto appena detto.<sup>(12)</sup> Di conseguenza, *qualsiasi sia l'anello commutativo unitario non nullo  $A$ , in  $A[x]$  esistono elementi non invertibili e diversi dallo zero, quindi  $A[x]$  non è un campo.*<sup>(13)</sup>

#### 4. DIVISIONE CON RESTO TRA POLINOMI

Se  $f$  e  $g$  sono due polinomi su un anello commutativo unitario  $A$ , con  $g \neq 0_A$ , diciamo che in  $A[x]$  è possibile effettuare la divisione di  $f$  (il *dividendo*) per  $g$  (il *divisore*) se e solo se esistono  $q, r \in A[x]$  (che chiamiamo rispettivamente *quoziente* e *resto*) tali che  $f = gq + r$  e  $\nu r < \nu g$ .

Un'osservazione banale è che se, nella situazione appena descritta,  $\nu f < \nu g$  allora è sicuramente possibile effettuare la divisione di  $f$  per  $g$ : basta porre  $q = 0$  e  $r = f$ . Il teorema che segue garantisce non solo la possibilità di effettuare la divisione, ma anche l'unicità di quoziente e resto in un caso importante.

**Teorema 6.** Siano  $A$  un anello commutativo unitario e  $f, g \in A[x]$ . Supponiamo  $\text{cd } g \in \mathcal{U}(A)$ . Allora esiste una ed una sola coppia  $(q, r) \in A[x] \times A[x]$  tale che  $f = gq + r$  e  $\nu r < \nu g$ .

*Dimostrazione.* Iniziamo a provare l'esistenza di  $(q, r)$ . Come appena osservato, se  $\nu f < \nu g$  una coppia con le proprietà richieste si ottiene ponendo  $q = 0$  e  $r = f$ . Possiamo allora supporre  $n := \nu f \geq m := \nu g$ ; osserviamo che l'ipotesi su  $\text{cd } g$  garantisce  $\text{cd } g \neq 0_A$  e quindi  $n, m \in \mathbb{N}$ . Ragioniamo per induzione su  $n$ , quindi supponiamo che, per ogni  $h \in A[x]$  tale che  $\nu h < n$ , sia possibile effettuare la divisione di  $h$  per  $g$ . Siano  $a = \text{cd } f$  e  $b = \text{cd } g$ . Consideriamo il polinomio  $k = ab^{-1}x^{n-m}g$ . È chiaro che per  $ab^{-1}x^{n-m}$  e  $g$  vale la regola di addizione dei gradi, in quanto il prodotto dei coefficienti direttori di questi due polinomi è  $(ab^{-1})\text{cd}(g) = ab^{-1} \cdot b = a \neq 0_A$  (vedi [Proposizione 3](#)). Dunque  $\nu k = \nu(ab^{-1}x^{n-m}) + \nu g = (n - m) + m = n = \nu f$  e  $\text{cd } k = a = \text{cd } f$ . Allora  $f$  e  $k$  hanno lo stesso grado,  $n$ , e lo stesso coefficiente direttore, quindi la [Proposizione 2](#) mostra che  $h := f - k$  ha grado minore di  $n$ . A questo punto l'ipotesi induttiva garantisce che è possibile effettuare la divisione di  $h$  per  $g$ , dunque esistono  $q_1, r_1 \in A[x]$  tali che  $h = gq_1 + r_1$  e  $\nu r_1 < \nu g$ . Ora,  $f = k + h = ab^{-1}x^{n-m}g + gq_1 + r_1 = g(ab^{-1}x^{n-m} + q_1) + r_1$ , quindi, la coppia  $(q, r)$ , definita ponendo  $q = ab^{-1}x^{n-m} + q_1$  e  $r = r_1$ , soddisfa le condizioni richieste. L'esistenza della coppia  $(q, r)$  è così provata.

Dobbiamo ora verificarne l'unicità. Siano  $(q, r)$  e  $(\bar{q}, \bar{r})$  due coppie con le proprietà richieste per quoziente e resto, dunque  $f = gq + r = g\bar{q} + \bar{r}$ , inoltre  $r, \bar{r} < m$ . Da  $gq + r = g\bar{q} + \bar{r}$  segue  $g(q - \bar{q}) = \bar{r} - r$ . Dalla [Proposizione 2](#) segue  $\nu(\bar{r} - r) < m$ . D'altra parte, poiché  $\text{cd } g$  è invertibile, quindi cancellabile, vale per  $g$  e  $q - \bar{q}$  la regola di addizione dei gradi, dunque  $\nu(g(q - \bar{q})) = m + \nu(q - \bar{q})$ . Abbiamo così  $m + \nu(q - \bar{q}) = \nu(\bar{r} - r) < m$ . Di conseguenza  $\nu(q - \bar{q}) = -\infty$ , quindi  $q - \bar{q} = 0_A$ , ovvero  $\bar{q} = q$  e, quindi, poiché  $\bar{r} - r = g(q - \bar{q}) = 0_A$ ,  $\bar{r} = r$ . L'unicità della coppia  $(q, r)$  è così dimostrata.  $\square$

Un caso molto importante è quello dei polinomi a coefficienti in un campo. Infatti, se  $A$  è un campo e  $0_A \neq g \in A[x]$  allora  $\text{cd } g$  è invertibile, come ogni elemento non nullo di  $A$ . Dunque, in questo caso, l'ipotesi  $\text{cd } g \in \mathcal{U}(A)$  nel [Teorema 6](#) può essere sostituita da  $g \neq 0_A$ . Abbiamo così:

**Corollario 7.** Siano  $K$  un campo e  $f, g \in K[x]$ . Se  $g \neq 0_K$  esiste una ed una sola coppia  $(q, r) \in K[x] \times K[x]$  tale che  $f = gq + r$  e  $\nu r < \nu g$ .

Notiamo che, con la notazione e nelle ipotesi del [Teorema 6](#),  $g$  divide  $f$  se e solo se il resto (unicamente determinato) della divisione di  $f$  per  $g$  è  $0_A$ .

Osserviamo poi che la dimostrazione del [Teorema 6](#) fornisce un algoritmo per il calcolo di quoziente e resto. Questo algoritmo non è altro che il procedimento comunemente insegnato anche nelle scuole per la divisione tra polinomi. Un esempio dovrebbe bastare a rendere questo punto chiaro. In  $\mathbb{Q}[x]$  consideriamo i polinomi  $f = 3x^5 + 3x^3 + x^2 - 1$  e  $g = 2x^3 + x + 3$ , e procediamo a dividere  $f$  per  $g$ . In accordo con le notazioni della dimostrazione del [Teorema 6](#), poniamo  $n = \nu f = 5$ ,  $m = \nu g = 3$ ,  $a = \text{cd } f = 3$  e  $b = \text{cd } g = 2$ . Abbiamo

<sup>(12)</sup>più in generale si può verificare, e si consiglia di farlo per esercizio, che lo stesso ragionamento mostra che se  $f$  è un polinomio non costante a coefficienti in un anello commutativo unitario  $A$  e  $\text{cd } f$  è cancellabile in  $A$ , allora  $f$  non è invertibile in  $A[x]$ .

<sup>(13)</sup>Abbiamo anche accennato in una nota precedente al fatto che se  $A$  è nullo, allora anche  $A[x]$  è nullo e quindi non è un campo. Dunque, *qualsiasi sia l'anello commutativo unitario non nullo  $A$ ,  $A[x]$  non è un campo.*

$ab^{-1}x^{n-m} = (3/2)x^2$  (che scriviamo sotto la linea al di sotto di  $g$ , perché sarà un addendo del quoziente) e  $k = ab^{-1}x^{n-m}g = 3x^5 + (3/2)x^3 + (9/2)x^2$ ; seguendo la procedura descritta nella dimostrazione del teorema calcoliamo  $h = f - k$ . Se si avesse  $\nu h < m$  allora la divisione sarebbe terminata:  $h$  sarebbe il resto, mentre il quoziente sarebbe  $ab^{-1}x^{n-m}$ .

The diagram illustrates the division of two polynomials  $f$  and  $g$ . On the left,  $f$  is divided by  $g$  to yield a quotient  $q$  and a remainder  $r$ . The terms are labeled with their degrees:  $3x^5$ ,  $3x^3$ ,  $x^2$  for  $f$ , and  $2x^3$ ,  $x$ ,  $3$  for  $g$ . The intermediate terms  $3x^5 + (3/2)x^3 + (9/2)x^2$  and  $(3/2)x^3 - (7/2)x^2$  are shown with dashed lines indicating they are subtracted from the previous row. The remainder  $r = -(7/2)x^2 - (3/4)x - 13/4$  is highlighted in a red box. On the right, the division of  $g$  by  $q$  is shown, with the remainder  $r = -(7/2)x^2 - (3/4)x - 13/4$  also highlighted in a red box.

Nel nostro esempio si ha invece  $h = (3/2)x^3 - (7/2)x^2 - 1$ , quindi  $\nu h \geq m$  (in questo esempio,  $\nu h = m$ ). La divisione va allora continuata, ripetendo la procedura dopo aver sostituito  $h$  ad  $f$ : posto  $a_1 = cd h$  e  $n_1 = \nu h$  calcoliamo  $a_1 b^{-1}x^{n_1-m}$  (che scriviamo come secondo addendo del quoziente) e poi  $k_1 = a_1 b^{-1}x^{n_1-m}g$  e  $h_1 = h - k_1$ . Nel nostro esempio abbiamo  $a_1 = 3/2$  e  $n_1 = 3$ , otteniamo dunque  $a_1 b^{-1}x^{n_1-m} = 3/4$ ,  $k_1 = (3/2)x^3 + (3/4)x + 9/4$  e  $h_1 = -(7/2)x^2 - (3/4)x - 13/4$ . Poiché  $\nu h_1 < m$  la divisione è terminata:  $h_1$  è il resto, il quoziente è la somma  $q = ab^{-1}x^{n-m} + a_1 b^{-1}x^{n_1-m}$  dei suoi addendi calcolati fino a questo punto. In altri casi avremmo potuto avere ancora  $h_1 \neq 0$  e  $\nu h_1 \geq m$  e la divisione non sarebbe terminata qui, ma la procedura avrebbe dovuto essere ancora ripetuta, dopo aver sostituito  $h_1$  ad  $f$ , calcolando, come nei passi precedenti, un polinomio  $h_2$ , di grado minore di  $\nu h_1$ , ed iterando ancora il procedimento fino ad ottenere un polinomio di grado minore di  $m$ : il resto della divisione; nello stesso tempo questo procedimento fornisce il quoziente come somma degli addendi calcolati ad ogni iterazione.

Il fatto che, nell'anello dei polinomi su un campo sia sempre possibile fare la divisione per un polinomio non nullo rende possibile, in questo caso, eseguire l'*algoritmo euclideo* delle divisioni successive per la ricerca del massimo comun divisore, esattamente allo stesso modo che nell'anello degli interi. Se  $K$  è un campo e  $f, g \in K[x]$ , se  $g \neq 0_K$  dividiamo  $f$  per  $g$  ottenendo un quoziente  $q$  ed un resto  $r$ , se  $r \neq 0_K$  dividiamo  $g$  per  $r$  ottenendo un resto  $r_1$ , se  $r_1 \neq 0_K$  dividiamo  $r$  per  $r_1$ ; se il resto  $r_2$  così ottenuto non è zero dividiamo  $r_1$  per  $r_2$  e così via. Questo procedimento termina dopo un numero finito di passi perché i successivi resti, finché sono diversi da  $0_K$ , hanno gradi strettamente decrescenti:  $\nu g > \nu r > \nu r_1 > \nu r_2 > \dots (\geq 0)$ ; dunque questa successione non può essere infinita. Così come per l'algoritmo euclideo in  $\mathbb{Z}$  (ed esattamente per lo stesso motivo) l'ultimo resto non nullo è un massimo comun divisore tra  $f$  e  $g$ . E, sempre come per  $\mathbb{Z}$ , si può estendere l'algoritmo e dimostrare (costruttivamente) il teorema di Bézout per i polinomi su campi:

**Teorema 8** (Teorema di Bézout). *Sia  $K$  un campo e siano  $f, g \in K[x]$ . Sia  $d$  un massimo comun divisore in  $K[x]$  tra  $f$  e  $g$ . Allora  $\{uf + vg \mid u, v \in K[x]\}$  coincide con l'insieme dei multipli di  $d$  in  $K[x]$ .*

**Esempio 9.** In  $\mathbb{Q}[x]$  consideriamo i polinomi  $f = 2x^5 - x^3 + 2x^2 - 1$  e  $g = x^5 + x^4 + x^3 + x^2 + x + 1$ . Eseguiamo l'algoritmo euclideo per calcolare un massimo comun divisore tra  $f$  e  $g$ . La divisione di  $f$  per  $g$  fornisce quoziente 2 e resto  $r = -2x^4 - 3x^3 - 2x - 3$ , infatti  $f = 2 \cdot g + r$ . Dividendo  $g$  per  $r$  otteniamo poi  $g = (-(1/2)x + 1/4) \cdot r + ((7/4)x^3 + 7/4)$ , qui  $q_1 = -(1/2)x + 1/4$  è il quoziente e  $r_1 = (7/4)x^3 + 7/4 = (7/4)(x^3 + 1)$  è il resto. La divisione successiva fornisce resto nullo, infatti  $r = (-(8/7)x - 12/7)r_1$ . Quindi  $r_1$ , l'ultimo resto non nullo, è un massimo comun divisore tra  $f$  e  $g$ . La teoria generale della divisibilità in monoidi commutativi ci dice che l'insieme dei massimi comuni divisori tra  $f$  e  $g$  è l'insieme dei polinomi associati a  $r_1$ ; come vedremo nelle prossime sezioni questo è l'insieme di tutti i polinomi della forma  $cr_1$  dove  $c$  è un numero razionale diverso da zero; tra questi massimi comuni divisori ne esiste dunque esattamente uno monico, precisamente  $(4/7)r_1 = x^3 + 1$ .

Come stabilito dal teorema di Bézout, possiamo scrivere  $r_1$  nella forma  $uf + gv$  per opportuni  $u, v \in \mathbb{Q}[x]$ . Possiamo calcolare una tale coppia  $(u, v)$  (ma sappiamo che ne esistono infinite) in questo modo: da  $g = q_1 \cdot r + r_1$  segue  $r_1 = g - q_1r$ ; inoltre da  $f = 2g + r$  segue  $r = f - 2g$ . Sostituendo questa espressione per  $r$  nell'uguaglianza precedente abbiamo  $r_1 = g - q_1(f - 2g) = (-q_1)f + (1 + 2q_1)g$ . Dunque, ponendo  $u = -q_1 = (1/2)x - 1/4$  e  $v = 1 + 2q_1 = -x + 3/2$ , l'uguaglianza  $r_1 = uf + gv$  è soddisfatta.

È bene tenere presente che l'algoritmo euclideo non può essere sempre utilizzato (almeno, non senza modifiche) per polinomi su anelli che non siano campi, come, ad esempio, in  $\mathbb{Z}[x]$ . Questo perché in questo caso non è sempre possibile effettuare la divisione tra polinomi non nulli; ad esempio, in  $\mathbb{Z}[x]$  non si può dividere  $2x^4 - 1$  per  $3x^2 + 1$  (perché?). È possibile (ma non lo facciamo qui) verificare che *il teorema di Bézout non vale nell'anello  $\mathbb{Z}[x]$* , dunque è essenziale, nel suo enunciato, richiedere che  $K$  sia un campo.

## 5. APPLICAZIONI POLINOMIALI E RADICI

Sia  $f \in A[x]$ , dove  $A$  è un anello commutativo unitario. Ricordiamo che se  $f = \sum_{i=0}^n a_i x^i$  e  $c \in A$  con  $f(c) = 0$  si indica l'elemento  $\sum_{i=0}^n a_i c^i$  di  $A$ . L'applicazione

$$\tilde{f}: c \in A \mapsto f(c) \in A$$

si chiama *applicazione polinomiale* determinata da  $f$  in  $A$ . A differenza dell'omomorfismo di sostituzione, definito nella [Sezione 2](#), quest'applicazione non è, in generale, un omomorfismo. Osserviamo che se  $f \in A$  allora  $f(c) = f$  per ogni  $c \in A$ , quindi l'applicazione  $\tilde{f}$  è costante. È per questo motivo che gli elementi di  $A$  vengono chiamati polinomi costanti (ma, attenzione!, è possibile che l'applicazione polinomiale  $\tilde{f}$  sia costante anche in casi in cui il polinomio  $f$  non è costante; vedremo [più avanti](#) qualche esempio di questo tipo).

Sempre nelle stesse notazioni, diciamo che  $c$  è una *radice* di  $f$  se e solo se  $f(c) = 0_A$ .

**Lemma 10.** Siano  $A$  un anello commutativo unitario e  $f, g \in A[x]$ . Allora:

- (i) se, in  $A[x]$ ,  $f$  divide  $g$ , ogni radice di  $f$  in  $A$  è radice di  $g$ ;
- (ii) se, in  $A[x]$ ,  $f$  e  $g$  sono associati,  $f$  e  $g$  hanno le stesse radici in  $A$ ;
- (iii) se  $A$  è un dominio di integrità, allora le radici di  $fg$  in  $A$  sono tutti e soli gli elementi di  $A$  che sono radici di  $f$  o di  $g$ .

*Dimostrazione.* (i): se  $f|_{A[x]} g$ , esiste  $h \in A[x]$  tale che  $g = fh$ . Allora, applicando l'omomorfismo di sostituzione definito da  $g$ , abbiamo  $g(c) = f(c)h(c) = 0_A h(c) = 0_A$ , dunque  $c$  è radice di  $f$ .<sup>(14)</sup>

(ii) segue subito da (i): se  $f$  e  $g$  sono associati,  $f$  divide  $g$  e  $g$  divide  $f$ , quindi le radici di  $f$  sono radici di  $g$  e viceversa.

(iii): Per la (i), gli elementi di  $A$  che sono radici di  $f$  o di  $g$  sono radici anche di  $fg$ , multiplo di entrambi. Viceversa, se  $c$  è una radice in  $A$  di  $fg$ , allora  $0_A = (fg)(c) = f(c)g(c)$ . Poiché, per la [Proposizione 5](#),  $A[x]$  è un dominio di integrità, questo implica che uno tra  $f(c)$  e  $g(c)$  è  $0_A$ , quindi  $c$  è radice di uno tra  $f$  e  $g$ .  $\square$

Esistono algoritmi che utilizzano questo semplice risultato per calcolare in modo efficiente valori di applicazioni polinomiali:

**Teorema 11** (Teorema del resto). Sia  $A$  un anello commutativo unitario e siano  $f \in A[x]$  e  $c \in A$ . Allora  $f(c)$  è il resto della divisione di  $f$  per  $x - c$ .

*Dimostrazione.* La prima cosa da osservare è che si può certamente effettuare la divisione di  $f$  per  $x - c$ , perché quest'ultimo polinomio è monico, quindi il suo coefficiente direttore è invertibile. Effettuata questa divisione, otteniamo  $q, r \in A[x]$  tali che  $f = (x - c)q + r$  e  $\nu r < \nu(x - c) = 1$ . Quest'ultima condizione equivale a dire che  $r$  è un polinomio costante, quindi  $r(c) = r$ . Applichiamo l'omomorfismo di sostituzione:  $f(c) = ((x - c)q + r)(c) = (c - c)q(c) + r(c) = 0_A q(c) + r = r$ . È così provato che  $f(c) = r$ .  $\square$

Dal teorema del resto si ottiene immediatamente:

**Teorema 12** (Teorema di Ruffini). Sia  $A$  un anello commutativo unitario e siano  $f \in A[x]$  e  $c \in A$ . Allora  $c$  è una radice di  $f$  se e solo se  $x - c$  divide  $f$  in  $A[x]$ .

*Dimostrazione.* Per il teorema del resto,  $c$  è radice di  $f$  se e solo se il resto della divisione di  $f$  per  $x - c$  è zero, cioè se e solo se  $x - c$  divide  $f$ .  $\square$

Ad esempio, una conseguenza del teorema di Ruffini è:

**Corollario 13.** Sia  $A$  un anello commutativo unitario e siano  $f, g \in A[x]$  e  $c \in A$ . Supponiamo che  $f$  e  $g$  abbiano in  $A[x]$  un massimo comun divisore  $d$ . Allora le radici comuni a  $f$  e  $g$  in  $A$  sono tutte e solo le radici di  $d$  in  $A$ :  $\{c \in A \mid f(c) = 0_A = g(c)\} = \{c \in A \mid d(c) = 0_A\}$ .

*Dimostrazione.* Sia  $c \in A$ . Per il teorema di Ruffini, dire che  $c$  è radice di  $f$  e di  $g$  equivale a dire che  $x - c$  è un divisore comune ad  $f$  e  $g$ . Per la definizione di massimo comun divisore, ciò equivale a dire che  $x - c$  divide  $d$ , quindi, ancora per il teorema di Ruffini, a dire che  $c$  è radice di  $d$ .  $\square$

Per polinomi su domini di integrità vale una versione più generale del teorema di Ruffini:

**Teorema 14** (Teorema di Ruffini generalizzato). Sia  $A$  un dominio di integrità unitario e siano  $f \in A[x]$ ,  $n \in \mathbb{N}^*$  e  $c_1, c_2, \dots, c_n$  elementi di  $A$  a due a due distinti. Allora si ha che ciascuno degli elementi  $c_i$  è radice di  $f$  se e solo se  $\prod_{i=1}^n (x - c_i)$  divide  $f$  in  $A[x]$ .

*Dimostrazione.* Una delle due implicazioni è ovvia: se  $\prod_{i=1}^n (x - c_i)$  divide  $f$  allora ciascuno degli elementi  $c_i$  è radice di  $f$ , in quanto  $x - c_i$  divide  $f$ . Dimostriamo l'implicazione inversa per induzione su  $n$ . Supponiamo che gli elementi  $c_1, c_2, \dots, c_n$  siano tutti radici di  $f$ . Se  $n = 1$  allora  $\prod_{i=1}^n (x - c_i) = x - c_1$  divide  $f$  per il teorema di Ruffini. Supponiamo allora  $n > 1$  e, come ipotesi di induzione, che l'enunciato valga per insiemi di  $n - 1$  elementi (distinti) di  $A$  ed arbitrari polinomi in  $A[x]$ . Poiché  $f(c_n) = 0_A$ , per il teorema di Ruffini esiste  $q \in A[x]$  tale che  $f = (x - c_n)q$ . Sia ora  $i$  un intero tale che  $1 \leq i < n$ . Poiché  $c_i$  è radice di  $f$  e  $A$  è un dominio di integrità, segue dal [Lemma 10](#) (iii) che  $c_i$  è radice di uno tra  $x - c_n$  e  $q$ . Ma  $c_i \neq c_n$ , per ipotesi, dunque  $(x - c_n)(c_i) = c_i - c_n \neq 0_A$ ; allora  $c_i$  non è radice di  $x - c_n$  e così  $c_i$  è radice di  $q$ . Dunque ciascuno degli elementi  $c_1, c_2, \dots, c_{n-1}$  è radice di  $q$ . Possiamo allora applicare l'ipotesi di induzione e concludere che  $\prod_{i=1}^{n-1} (x - c_i)$  divide  $q$ , quindi esiste  $h \in A[x]$  tale che  $q = h \prod_{i=1}^{n-1} (x - c_i)$ . Allora  $f = q \cdot (x - c_n) = h(\prod_{i=1}^{n-1} (x - c_i)) \cdot (x - c_n) = h \prod_{i=1}^n (x - c_i)$ . Pertanto  $\prod_{i=1}^n (x - c_i)$  divide  $f$ ; la dimostrazione è così completa.  $\square$

<sup>(14)</sup>in alternativa, si potrebbe dedurre la (i) dal teorema di Ruffini. Come?

Il teorema di Ruffini generalizzato ha due importantissime conseguenze. La prima è una limitazione al numero di radici che un polinomio non nullo su un dominio di integrità può avere.

**Teorema 15.** *Sia  $A$  un dominio di integrità unitario e sia  $0_A \neq f \in A[x]$ . Allora il numero delle radici di  $f$  in  $A$  non supera  $\nu f$ .*

*Dimostrazione.* Se  $f$  ha esattamente  $n$  radici, siano esse  $c_1, c_2, \dots, c_n$ , allora  $f$  è multiplo di  $g := \prod_{i=1}^n (x - c_i)$ , per il teorema di Ruffini generalizzato, quindi  $f = gq$  per opportuno  $q \in A[x]$ . Essendo  $f \neq 0_A$  si ha anche  $q \neq 0_A$ . Ma  $\nu g = n$  e per  $g$  e  $q$  vale la regola di addizione dei gradi (perché  $A$  è un dominio di integrità, o, in alternativa, perché  $g$  è monico, quindi ha coefficiente direttore invertibile). Quindi  $\nu f = \nu g + \nu q = n + \nu q \geq n$ .  $\square$

Sia per il teorema di Ruffini generalizzato che per il [Teorema 15](#) è essenziale l'ipotesi che  $A$  sia un dominio di integrità. Consideriamo, ad esempio, il polinomio  $f = \bar{2}x \in \mathbb{Z}_6[x]$ . Sia  $[0]_6$  che  $[3]_6$  sono radici di  $f$ , quindi  $f$  ha più radici in  $\mathbb{Z}_6$  di quanto sia il suo grado, che è 1. Come imposto dal teorema di Ruffini sia  $x = x - [0]_6$  che  $x - [3]_6$  dividono  $f$  (infatti  $f = x \cdot [2]_6 = (x - [3]_6) \cdot [2]_6$ ), ma  $x(x - [3]_6)$  non divide  $f$ , quindi la conclusione del teorema di Ruffini generalizzato non vale per  $f$ .

L'altra conseguenza del teorema di Ruffini generalizzato riguarda le applicazioni polinomiali e ci dice che nel caso dei domini di integrità infiniti ogni polinomio è identificato univocamente dalla sua applicazione polinomiale.

**Teorema 16** (Principio di identità dei polinomi). *Sia  $A$  un dominio di integrità infinito. Allora, per ogni  $f, g \in A[x]$  si ha:  $\tilde{f} = \tilde{g} \iff f = g$ .*

*Dimostrazione.* Ovviamente  $\tilde{f} = \tilde{g}$  se  $f = g$ . Supponiamo, viceversa,  $\tilde{f} = \tilde{g}$ . Allora  $f(c) = g(c)$  per ogni  $c \in A$ . Sia  $h = f - g$ . Allora per ogni  $c \in A$  abbiamo  $h(c) = (f - g)(c) = f(c) - g(c) = 0_A$ , vale a dire: ogni elemento di  $A$  è radice di  $h$ . Dunque  $h$  ha un numero infinito di radici. Ma il [Teorema 15](#) assicura che se  $h \neq 0_A$  allora il numero delle radici di  $h$  non supera  $\nu h$ , quindi è finito. Di conseguenza deve essere  $h = 0_A$ , vale a dire:  $f = g$ .  $\square$

È a causa del principio di identità dei polinomi che in alcuni casi vengono identificati i polinomi con le applicazioni polinomiali. Ad esempio, nei corsi di analisi matematica si definiscono i polinomi come particolari funzioni da  $\mathbb{R}$  a  $\mathbb{R}$ , quelle che per noi sono le applicazioni polinomiali definite dai polinomi in  $\mathbb{R}[x]$ . Questo è lecito perché, essendo  $\mathbb{R}$  un campo (quindi un dominio di integrità) infinito, il principio di identità dei polinomi assicura che i polinomi in  $\mathbb{R}[x]$  corrispondono esattamente alle loro applicazioni polinomiali (in corsi di analisi più avanzati i polinomi sono definiti con riferimento al campo complesso, anziché a quello reale; il discorso è analogo: anche per il campo complesso vale il principio di identità dei polinomi). D'altra parte, non è lecito identificare polinomi ed applicazioni polinomiali in contesti in cui non valga il principio di identità dei polinomi, cioè quando l'anello  $A$  considerato sia finito oppure non sia integro.

Nel caso degli anelli finiti è certo che il principio di identità dei polinomi non può valere. Infatti, se  $A$  è un anello commutativo unitario finito, il numero delle applicazioni da  $A$  ad  $A$ , e quindi il numero delle applicazioni polinomiali in  $A$ , è finito, mentre  $A[x]$  è comunque infinito. Dunque, in questo caso, è impossibile che ci sia una corrispondenza biunivoca tra polinomi e applicazioni polinomiali (cioè che il principio di identità dei polinomi afferma è che, se  $A$  è un dominio di integrità infinito, l'applicazione  $f \in A[x] \mapsto \tilde{f} \in \text{Map}(A, A)$  è iniettiva; ciò è impossibile nel caso che stiamo considerando ora, in cui il dominio  $A[x]$  è infinito ma il codominio  $\text{Map}(A, A)$  è finito). Possiamo fare esempi più esplicativi: il polinomio  $x(x - \bar{1})(x - \bar{2}) = x^3 - x \in \mathbb{Z}_3[x]$  ha tutti gli elementi del campo  $\mathbb{Z}_3$  come radici, questo significa che  $\tilde{f}$  è l'applicazione costante  $c \in \mathbb{Z}_3 \mapsto \bar{0} \in \mathbb{Z}_3$ , ma allora  $\tilde{f}$  coincide con l'applicazione polinomiale  $\tilde{0}_{\mathbb{Z}_3}$  definita dal polinomio nullo, pur essendo  $f \neq 0_{\mathbb{Z}_3}$ . Più in generale, se se  $F$  è un campo finito, di cardinalità  $q$ , il polinomio  $f = \prod_{c \in F} (x - c)$  ha grado  $q$  ed ha tutti gli elementi di  $F$  come radici, quindi  $\tilde{f} = \tilde{0}_F$ . È possibile dimostrare che due polinomi in  $F[x]$  definiscono la stessa applicazione polinomiale se e solo se la loro differenza è un multiplo di questo polinomio  $f$ .

Anche nel caso degli anelli infiniti, che però non siano integri, il principio di identità dei polinomi può non valere. Ad esempio, se  $A$  è un anello booleano e  $f = x^2 - x \in A[x]$  allora, poiché ogni elemento  $c$  di  $A$  è idempotente e quindi verifica  $c^2 - c = 0_A$ , ovvero  $f(c) = 0_A$ , tutti gli elementi di  $A$  sono radici di  $f$ . Allora  $\tilde{f} = \tilde{0}_A$ , anche se  $f \neq 0_A$ . Nel caso in cui  $A$  sia infinito,  $f$  è un esempio di polinomio di secondo grado con infinite radici.

## 6. FATTORIZZAZIONE

Ricordiamo che un monoide commutativo cancellativo (cioè ad elementi tutti cancellabili) si dice *fattoriale* se e solo se ogni suo elemento non invertibile è prodotto di elementi irriducibili e tali decomposizioni in irriducibili sono essenzialmente uniche.<sup>(15)</sup> Se  $A$  è un dominio di integrità unitario, allora  $A^\# := A \setminus \{0_A\}$  è chiuso rispetto

(15)quest'ultima frase vuol dire che se  $r, s \in \mathbb{N}^*$  e  $p_1, p_2, \dots, p_r, q_1, q_2, \dots, q_s$  sono elementi irriducibili tali che  $p_1 p_2 \cdots p_r = q_1 q_2 \cdots q_s$  allora  $r = s$  ed esiste una permutazione  $\sigma \in \mathbb{S}_r$  tale che, per ogni  $i \in \{1, 2, \dots, r\}$  gli elementi  $p_i$  e  $q_{i\sigma}$  siano associati. Detto in modo più facile (ma più ambiguo): il numero dei fattori nei prodotti  $p_1 p_2 \cdots p_r$  e  $q_1 q_2 \cdots q_s$  è lo stesso, ed i fattori nel secondo prodotto possono essere riordinati in modo che fattori corrispondenti nei due prodotti ( $p_1$  con  $q_1$ ,  $p_2$  con  $q_2$  etc.) siano associati tra loro. Già che ci siamo, ricordiamo anche che un elemento  $p$  si dice *irriducibile* se e solo se non è invertibile ed i suoi unici divisori sono quelli banali, cioè gli elementi invertibili e quelli associati a  $p$ . Un elemento non invertibile né irriducibile è invece *riducibile*. Due elementi  $a$  e  $b$  di un monoide commutativo  $S$  sono, per definizione, *associati* se e solo se  $a$  divide  $b$  e  $b$  divide  $a$  (in  $S$ ); se poi  $S$  è anche un cancellativo si dimostra che  $a$  e  $b$  sono associati in  $S$  se e solo se  $b = au$  per un opportuno elemento invertibile  $u$  di  $S$ . In ogni caso, la relazione 'essere elementi associati' è di equivalenza in  $S$  ed elementi tra loro associati hanno esattamente gli stessi divisori e gli stessi multipli.

alla moltiplicazione (questa affermazione è esattamente la legge di annullamento del prodotto: il prodotto tra due elementi di  $A$  diversi da zero è diverso da zero), quindi, con la moltiplicazione indotta da quella dell'anello,  $A^\#$  è un monoide, cancellativo perché sono cancellabili in  $A$  tutti i suoi elementi. Si dice che  $A$  è un *anello fattoriale* se e solo se questo monoide  $A^\#$  è fattoriale. Il motivo per cui questa nozione è importante nello studio degli anelli di polinomi è il seguente teorema, che non dimostreremo:

**Teorema 17.** Se  $A$  è un anello fattoriale allora  $A[x]$  è fattoriale.

Ora, sono certamente fattoriali i campi ed è fattoriale, per il Teorema Fondamentale dell'Aritmetica, l'anello  $\mathbb{Z}$  degli interi. Quindi è fattoriale  $\mathbb{Z}[x]$  e, per ogni campo  $K$ , anche  $K[x]$  (che, come già osservato, non può essere un campo). Dunque, sia per i polinomi a coefficienti in  $\mathbb{Z}$  che per quelli a coefficienti in un campo vale un teorema di fattorizzazione essenzialmente unica in prodotto di polinomi irriducibili: ogni polinomio non invertibile è prodotto di polinomi irriducibili e tale fattorizzazione è unica a meno dell'ordine dei fattori e della sostituzione di alcuni fattori con polinomi associati.

Nell'ipotesi che  $A$  sia fattoriale, una delle conseguenze del fatto che  $A[x]$  è fattoriale è che (in analogia con ciò che accade in  $\mathbb{Z}$ ), nota una fattorizzazione in prodotto di irriducibili di un polinomio  $f$  è facile determinare l'insieme dei divisori di  $f$ . Diamo un rapido cenno, tutto funziona come nell'aritmetica in  $\mathbb{Z}$ : posto  $f = p_1^{\lambda_1} p_2^{\lambda_2} \cdots p_n^{\lambda_n}$ , dove ciascuno dei polinomi  $p_i$  è irriducibile, se  $i \neq j$  allora  $p_i$  e  $p_j$  non sono associati e  $\lambda_i \in \mathbb{N}$  per ogni  $i$ , i divisori di  $f$  sono tutti e soli i polinomi della forma  $p_1^{\sigma_1} p_2^{\sigma_2} \cdots p_n^{\sigma_n}$  ed i loro associati, dove, per ogni  $i$ , valga  $\sigma_i \in \mathbb{N}$  e  $\sigma_i \leq \lambda_i$ . Usando questa osservazione è possibile anche notare che, scelti comunque  $f, g \in A[x]$ , esistono un massimo comun divisore  $d$  ed un minimo comune multiplo  $m$  tra  $f$  e  $g$ , e  $dm$  è associato a  $fg$ . Infatti, escluso il caso banale in cui uno tra  $f$  e  $g$  è il polinomio nullo,  $f$  e  $g$  si possono fattorizzare nella forma  $f = up_1^{\lambda_1} p_2^{\lambda_2} \cdots p_n^{\lambda_n}$  e  $g = vp_1^{\mu_1} p_2^{\mu_2} \cdots p_n^{\mu_n}$ , dove, come sopra i  $p_i$  sono irriducibili a due a due non associati,  $u, v \in \mathcal{U}(A[x]) = \mathcal{U}(A)$  e  $\lambda_i, \mu_i \in \mathbb{N}$  per ogni  $i$  (notare che non è escluso che alcuni dei  $\lambda_i$  o  $\mu_i$  siano 0). Si verifica allora che, ponendo  $\sigma_i = \min\{\lambda_i, \mu_i\}$  e  $\tau_i = \max\{\lambda_i, \mu_i\}$  per ogni  $i$ , si ha che  $d := p_1^{\sigma_1} p_2^{\sigma_2} \cdots p_n^{\sigma_n}$  è un massimo comun divisore e  $m := p_1^{\tau_1} p_2^{\tau_2} \cdots p_n^{\tau_n}$  è un minimo comune multiplo tra  $f$  e  $g$ ; inoltre  $dm$  è associato a  $fg$  perché  $\sigma_i + \tau_i = \lambda_i + \mu_i$  per ogni  $i$ , quindi  $fg = uvdm$ .

Nella pratica, è spesso molto più difficile calcolare una fattorizzazione in irriducibili di un polinomio che eseguire l'algoritmo euclideo, quindi risulta in genere conveniente questo secondo metodo quando si ricerca un massimo comun divisore tra due polinomi. È bene però notare che la discussione appena svolta mostra che anche nei casi in cui l'algoritmo euclideo non si può eseguire, come ad esempio in  $\mathbb{Z}[x]$ , la fattorialità garantisce comunque l'esistenza di un massimo comun divisore e di un minimo comune multiplo tra due polinomi.

Passiamo ora a discutere in maggior dettaglio le fattorizzazioni in polinomi irriducibili. Iniziamo a stabilire: quando è che due polinomi sono associati? Se  $A$  è un dominio di integrità unitario e  $f \in A[x]$ , i polinomi associati ad  $f$  sono tutti e soli quelli della forma  $uf$ , dove  $u$  è un polinomio invertibile in  $A[x]$ .<sup>(16)</sup> Come sappiamo ([Proposizione 5](#)),  $\mathcal{U}(A[x]) = \mathcal{U}(A)$ , quindi i polinomi associati ad  $f$  sono tutti e soli i polinomi della forma  $uf$ , dove  $u$  è un invertibile di  $A$ . In questo caso, quindi, *polinomi (non nulli) associati hanno necessariamente lo stesso grado*. Ad esempio, poiché  $\mathcal{U}(\mathbb{Z}) = \{1, -1\}$ , gli associati in  $\mathbb{Z}[x]$  di un  $f \in \mathbb{Z}[x]$  sono  $f$  stesso (cioè  $1f$ ) e  $-f$  (cioè  $(-1)f$ ). Se invece  $K$  è un campo  $\mathcal{U}(K[x]) = \mathcal{U}(K) = K^\# := K \setminus \{0_K\}$ , quindi, l'insieme di tutti i polinomi associati ad un  $f \in K[x] \setminus \{0_K\}$  è  $\{cf \mid 0_K \neq c \in K\}$ . Se  $a = cd f$  si ha  $cd(cf) = ca$  per ogni  $c \in K^\#$ . Allora, qualunque sia  $k \in K^\#$ , il nostro polinomio  $f$  ha esattamente un associato con coefficiente direttore  $k$ , precisamente  $(ka^{-1})f$  (infatti, per ogni  $c \in K^\#$ , abbiamo che  $cd(cf) = ca = k$  se e solo se  $c = ka^{-1}$ ). Il caso più importante è quello in cui scegliamo  $k = 1_K$ . In questo caso ciò che otteniamo è che  $f$  ha un unico associato con coefficiente direttore  $1_K$ , ovvero monico, precisamente  $a^{-1}f$ . Otteniamo così:

**Proposizione 18.** Sia  $K$  un campo. In ogni classe di elementi associati di polinomi non nulli in  $K[x]$  esiste uno ed un solo polinomio monico.

Ci riferiamo a questo polinomio come al *rappresentante monico* della classe. A titolo di esempio, in  $\mathbb{Q}[x]$  il polinomio monico associato a  $f := 3x^2 + x - 6$  è  $(1/3)f = x^2 + (1/3)x - 2$ ; ma anche  $-6x^2 - 2x + 12$  e  $(1/100)x^2 + (1/300)x - 1/50$  (e infiniti altri polinomi, tutti quelli della forma  $cf$ , dove  $0 \neq c \in \mathbb{Q}$ ) sono associati a  $f$ . L'esistenza di un unico rappresentante monico in ogni classe di polinomi non nulli associati permette di esprimere le fattorizzazioni in irriducibili dei polinomi a coefficienti in un campo in una forma spesso più conveniente:

**Proposizione 19.** Sia  $K$  un campo. Allora ogni polinomio non nullo in  $K[x]$  è prodotto di un elemento di  $K$  e di polinomi monici irriducibili in  $K[x]$ . Tale fattorizzazione è unica a meno dell'ordine dei fattori.

*Dimostrazione.* L'unicità della fattorizzazione segue dal fatto che  $K[x]$  è fattoriale e dal fatto che ogni classe di polinomi associati non nulli contiene un solo rappresentante monico. L'esistenza della decomposizione è ovvia nel caso dei polinomi costanti, va provata per polinomi non costanti. Sia, allora,  $f \in K[x] \setminus K$ . Sia  $f = p_1 p_2 \cdots p_n$  una fattorizzazione di  $f$  in prodotto di polinomi irriducibili. Per ogni  $i \in \{1, 2, \dots, n\}$  sia  $a_i = cd(p_i)$ ; allora  $p_i = a_i q_i$ , dove  $q_i = a_i^{-1} p_i$  è associato a  $p_i$  (quindi è irriducibile) ed è monico. Posto  $a = a_1 a_2 \cdots a_n$  abbiamo  $f = a q_1 q_2 \cdots q_n$ ; questa è la decomposizione cercata.  $\square$

Si noti che, nella fattorizzazione appena descritta,  $a = cd f$ .

<sup>(16)</sup>questo è vero se  $f \neq 0_A$ , perché in questo caso  $f$  è cancellabile, ma è anche banalmente vero se  $f = 0_A$ .

Ci vogliamo ora occupare di descrivere, per quanto possibile, la proprietà di essere o meno irriducibile per un polinomio a coefficienti in un campo. Vedremo in che modo questo proprietà è collegata alla presenza di radici. Iniziamo con una importante caratterizzazione, che, quando la trattazione è limitata ad anelli di polinomi su campi, è talvolta utilizzata per definire la nozione di polinomio irriducibile. Va tenuto ben presente che, come vedremo, questo teorema non si applica a polinomi su anelli che non siano campi.

**Teorema 20.** *Siano  $K$  un campo e  $f \in K[x]$ . Se  $n = \nu f$  allora  $f$  è irriducibile in  $K[x]$  se e solo se  $n > 0$  e vale una delle due proprietà equivalenti:*

- (a) *non esistono  $g, h \in K[x]$  tali che  $f = gh$  e sia  $g$  che  $h$  abbiano grado minore di  $n$ ;*
- (b) *non esistono  $g, h \in K[x]$  tali che  $f = gh$  e sia  $g$  che  $h$  abbiano grado maggiore di  $0$ .*

*Dimostrazione.*  $f$  è irriducibile se e solo se non è invertibile e non ha divisori se non quelli banali. Se  $f \in K$ , allora o  $f$  è il polinomio nullo, che non è irriducibile perché gli elementi non nulli di  $K[x]$  sono suoi divisori banali, oppure è un elemento di  $\mathcal{U}(K) = K \setminus \{0_K\}$ , ed è quindi invertibile in  $K[x]$ , per la [Proposizione 5](#). Dunque, se  $f$  è irriducibile, sicuramente  $f \notin K$ , vale a dire:  $n > 0$ . Supponiamo dunque  $n > 0$ . Osserviamo che, se  $g, h \in K[x]$  e  $f = gh$ , per la regola di addizione dei gradi (che vale perché  $K$  è un campo) si ha  $\nu g + \nu h = \nu f = n$ , quindi  $(\nu g < n \wedge \nu h < n) \iff (\nu g > 0 \wedge \nu h > 0)$ , vale a dire: (a) e (b) sono equivalenti. Se  $f$  è irriducibile, scelti comunque  $g, h \in K[x]$  tali che  $f = gh$ , allora  $g$  è un divisore di  $f$ , quindi un divisore banale perché  $f$  è irriducibile. Allora o  $g$  è invertibile, nel qual caso  $g \in K \setminus \{0_K\}$  e  $\nu g = 0$ , oppure  $g$  è associato a  $f$ , nel qual caso  $\nu g = \nu f = n$ . Ciò mostra che, se  $f$  è irriducibile, sono verificate (a) e (b). Se, invece,  $f$  non è irriducibile,  $f$  ha un divisore non banale  $g$ ; allora  $g \neq 0_K$  (altrimenti  $f = 0_K$ ) e  $g$  non è invertibile, quindi  $\nu g > 0$ , ed esiste  $h \in K[x]$  tale che  $f = gh$ . Ovviamente  $h \neq 0_K$ , e  $h$  non è invertibile perché  $g$  non è associato ad  $f$ , quindi abbiamo anche  $\nu h > 0$ . In questo caso, dunque, non vale (b), e quindi neanche (a).  $\square$

Se  $K$  è un campo, ogni polinomio di primo grado  $ax + b \in K[x]$  ha una radice in  $K$  (precisamente  $-a^{-1}b$ : essendo il polinomio di grado 1 si ha  $a \neq 0_K$  quindi ha senso considerare  $a^{-1}$ ), dunque un polinomio  $f$  che sia divisibile per un polinomio di primo grado ha almeno una radice in  $K$ , per il [Lemma 10](#). Viceversa, se  $f$  ha una radice allora  $f$  ha un divisore di primo grado, per il teorema di Ruffini. Dunque:

**Proposizione 21.** *Sia  $K$  un campo e sia  $f \in K[x]$ . Allora  $f$  ha radici in  $K$  se e solo se ha almeno un divisore di primo grado in  $K[x]$ .*

Siccome una delle due implicazioni, quella stabilita dal teorema di Ruffini, vale per polinomi su anelli commutativi unitari qualsiasi, possiamo anche osservare:

**Proposizione 22.** *Sia  $A$  un dominio di integrità unitario e sia  $0_A \neq f \in A[x]$ . Se  $\nu f > 1$  e  $f$  ha radici in  $A$ , allora  $f$  è riducibile in  $A[x]$ .*

*Dimostrazione.* Per il teorema di Ruffini,  $f$  ha un divisore  $h$  di primo grado. Allora  $h$  non è invertibile (per la [Proposizione 5](#)) e poiché, come già osservato, due polinomi in  $A[x]$  che siano associati devono avere lo stesso grado, mentre  $\nu h = 1 < \nu f$ , allora  $h$  non è associato a  $f$ . Pertanto  $h$  è un divisore non banale di  $f$ , quindi  $f$  non è irriducibile.  $\square$

In un caso molto particolare, ma importante, vale anche il viceversa:

**Proposizione 23.** *Siano  $K$  un campo e  $f$  un polinomio non nullo in  $K[x]$  di grado 2 o 3. Allora  $f$  è irriducibile in  $K[x]$  se e solo se è privo di radici in  $K$ .*

*Dimostrazione.* Poiché  $\nu f > 0$ , certamente  $f$  non è invertibile. Se  $f$  è irriducibile allora è privo di radici, per la [Proposizione 22](#). Viceversa, se  $f$  è riducibile allora per la [Teorema 20](#) dobbiamo avere  $f = gh$  per opportuni  $g, h \in K[x]$  tali che  $\nu g, \nu h < \nu f$ , e naturalmente  $\nu g + \nu h = \nu f$ . Se  $\nu f = 2$  abbiamo una sola possibilità:  $\nu g = \nu h = 1$ ; se  $\nu f = 3$  abbiamo invece due casi possibili:  $\nu g = 1$  e  $\nu h = 2$  oppure, viceversa,  $\nu g = 2$  e  $\nu h = 1$ . In tutti e tre i casi, comunque,  $f$  ha un divisore di grado 1, quindi una radice. Con questo l'enunciato è dimostrato.  $\square$

Possiamo schematizzare come segue le informazioni ottenute sulle proprietà di un polinomio a coefficienti in un campo di essere o meno irriducibile ed di avere o meno radici.

Se  $K$  è un campo e  $0_K \neq f \in K[x]$ , posto  $n = \nu f$  si ha:

$n = 0$	$\implies f$ è invertibile e privo di radici
$n = 1$	$\implies f$ è irriducibile ed ha una radice
$n \in \{2, 3\}$	$(f$ è irriducibile $\iff f$ non ha radici)
$n > 3$	$(f$ è irriducibile $\implies f$ non ha radici)

(Ovviamente qui per ‘irriducibile’ si intende ‘irriducibile in  $K[x]$ ’ e per ‘radice’ si intende ‘radice in  $K$ ’).

Osserviamo che l’implicazione all’ultimo rigo di questa tabella, in generale, non si inverte. Ad esempio, un polinomio di grado 4 può essere il prodotto di due polinomi irriducibili di grado 2; in questo caso il polinomio è irriducibile (ovviamente ...) ma privo di radici (perché privo di divisori di primo grado; oppure per questo motivo:

una radice dovrebbe necessariamente essere radice di uno dei fattori di grado due, ma essendo irriducibili questi sono privi di radici). Un esempio di questo tipo è il polinomio  $(x^2 + 1)(x^2 + 2)$  in  $\mathbb{Q}[x]$ .

Non va poi dimenticato che tutti questi risultati valgono nel caso dei polinomi a coefficienti in un campo, ma (ad eccezione della [Proposizione 22](#)) non in casi più generali. Ad esempio, in  $\mathbb{Z}[x]$  il polinomio (costante) 2 è irriducibile (non invertibile!) in  $\mathbb{Z}[x]$ , pur avendo grado 0; il polinomio  $2x$ , che è irriducibile in  $\mathbb{Q}[x]$  perché  $\mathbb{Q}$  è un campo e  $\nu(2x) = 1$ , è invece riducibile in  $\mathbb{Z}[x]$ , perché è diviso da 2 che, in  $\mathbb{Z}[x]$  non è invertibile né associato a  $2x$ , quindi è un divisore non banale di  $2x$ . Come si vede, la differenza sta nel fatto che 2 è invertibile in  $\mathbb{Q}[x]$  ma non in  $\mathbb{Z}[x]$ . Inoltre, in  $\mathbb{Z}[x]$  il polinomio di primo grado  $2x + 1$  è privo di radici, quindi anche la [Proposizione 21](#) non vale per arbitrari polinomi su  $\mathbb{Z}$ .

## 7. METODI ED ESEMPI DI FATTORIZZAZIONE PER POLINOMI SU UN CAMPO

Supponiamo di voler fattorizzare un polinomio (in un fissato anello di polinomi) in prodotto di polinomi irriducibili. Per farlo abbiamo bisogno:

- di saper trovare divisori non banali del polinomio dato, se ne esistono;
- di saper riconoscere quali tra questi divisori sono irriducibili.

Limitiamoci al caso dei polinomi su un campo. Usando la tabella nella sezione precedente, sappiamo, in linea di massima, rispondere al secondo punto nel caso di divisori di grado minore di quattro. I polinomi di grado uno sono sempre irriducibili, quelli di grado due o tre lo sono se e solo se sono privi di radici. In due casi notevoli queste informazioni sono addirittura più di quanto non sia necessario. Infatti valgono questi teoremi (che non dimostriamo) per polinomi in  $\mathbb{C}[x]$  ed in  $\mathbb{R}[x]$  (come di consueto,  $\mathbb{C}$  indica il campo dei numeri complessi ed  $\mathbb{R}$  il campo dei numeri reali).

**Teorema 24.** *Ogni polinomio non costante in  $\mathbb{C}[x]$  ha qualche radice in  $\mathbb{C}$ . Di conseguenza, gli unici polinomi irriducibili in  $\mathbb{C}[x]$  sono quelli di grado uno.*

**Teorema 25.** *Ogni polinomio irriducibile in  $\mathbb{R}[x]$  ha grado minore di 3.*

Dunque, i polinomi irriducibili in  $\mathbb{R}[x]$  sono precisamente quelli di grado 1 e quelli di grado 2 privi di radici. Come è noto dalle scuole superiori, un polinomio  $ax^2 + bx + c \in \mathbb{R}[x]$  di grado 2 ha radici in  $\mathbb{R}$  se e solo se  $b^2 - 4ac \geq 0$ . Dunque, è molto facile riconoscere se un polinomio in  $\mathbb{C}[x]$  o in  $\mathbb{R}[x]$  è irriducibile. A proposito dei polinomi in  $\mathbb{R}[x]$  vale anche questo risultato, che si può provare con metodi elementari dell'analisi (è una conseguenza del teorema di Bolzano):

**Teorema 26.** *Ogni polinomio di grado dispari in  $\mathbb{R}[x]$  ha qualche radice in  $\mathbb{R}$ .*

Osserviamo che quest'ultimo teorema di potrebbe anche dedurre dal precedente, se si supponesse di aver dimostrato quello. Infatti, se  $f$  è un polinomio di grado dispari in  $\mathbb{R}[x]$  e  $f = p_1 p_2 \cdots p_r$  è una sua fattorizzazione in prodotto di polinomi irriducibili in  $\mathbb{R}[x]$ , allora, dal momento che ciascuno dei polinomi  $p_i$  ha grado 1 o 2, ma non tutti possono avere grado 2, altrimenti  $\nu(f) = \sum_{i=1}^r \nu(p_i)$  sarebbe  $2r$ , che è pari, almeno uno dei fattori  $p_i$  deve avere grado 1, quindi  $f$  ha un divisore di primo grado e così ha una radice.

La situazione è molto più complessa (ed interessante) nel caso di polinomi in  $\mathbb{Q}[x]$ . Lo studio dei polinomi in  $\mathbb{Q}[x]$  si può ridurre al caso dei polinomi a coefficienti interi. Infatti, se  $f = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \in \mathbb{Q}[x]$ , ciascuno dei coefficienti  $a_i$  sarà una frazione, che possiamo scrivere come  $a_i = u_i/v_i$ , dove  $u_i, v_i \in \mathbb{Z}$  e  $v_i \neq 0$ . Se  $m$  è un multiplo comune a  $v_0, v_1, \dots, v_n$  e  $m \neq 0$  allora  $\bar{f} := mf \in \mathbb{Z}[x]$ ; poiché  $m \in \mathcal{U}(\mathbb{Q}[x])$ , inoltre,  $\bar{f}$  è associato a  $f$  in  $\mathbb{Q}[x]$ . Pertanto:

**Lemma 27.** *Ogni polinomio  $f \in \mathbb{Q}[x]$  è associato, in  $\mathbb{Q}[x]$ , ad un polinomio  $\bar{f} \in \mathbb{Z}[x]$ .*

Ora, polinomi tra loro associati hanno esattamente le stesse proprietà rispetto alla fattorizzazione. Ad esempio, i polinomi  $f$  e  $\bar{f}$  di questo lemma hanno gli stessi divisori in  $\mathbb{Q}[x]$ ,  $f$  è irriducibile in  $\mathbb{Q}[x]$  se e solo se lo è  $\bar{f}$ , inoltre  $f$  e  $\bar{f}$  hanno esattamente le stesse radici in  $\mathbb{Q}$  ([Lemma 10](#), ad esempio). In questo senso lo studio di  $\bar{f}$  equivale allo studio di  $f$ . Bisogna però fare attenzione, ci stiamo riferendo a  $\bar{f}$  riguardato come polinomio in  $\mathbb{Q}[x]$ . Detto diversamente, ci interessano le proprietà di fattorizzazione di  $\bar{f}$  in  $\mathbb{Q}[x]$ , non in  $\mathbb{Z}[x]$ . Come sappiamo già da esempi visti in precedenza, anche per polinomi in  $\mathbb{Z}[x]$  le proprietà di essere irriducibile in  $\mathbb{Z}[x]$  o di essere irriducibile in  $\mathbb{Q}[x]$  non sono equivalenti;  $\bar{f}$  potrebbe essere irriducibile in  $\mathbb{Q}[x]$  pur non essendolo in  $\mathbb{Z}[x]$ .

A differenza di quanto accade in  $\mathbb{C}[x]$  ed in  $\mathbb{R}[x]$ , esistono in  $\mathbb{Q}[x]$  polinomi irriducibili di grado arbitrariamente grande. Questo segue dal prossimo teorema, che fornisce un utile criterio sufficiente a dimostrare l'irriducibilità di alcuni polinomi.

**Teorema 28** (Criterio di irriducibilità di Eisenstein). *Sia  $f = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \in \mathbb{Z}[x]$ . Se esiste un primo  $p$  tale che:*

- (1)  $p$  divide  $a_0, a_1, \dots, a_{n-1}$ ,
- (2)  $p$  non divide  $a_n$ ,
- (3)  $p^2$  non divide  $a_0$ ,

allora  $f$  è irriducibile in  $\mathbb{Q}[x]$ .

Ad esempio, per ogni intero positivo  $n$  e per ogni primo  $p$ , il polinomio  $x^n - p$  è irriducibile in  $\mathbb{Q}[x]$ . Infatti possiamo applicare il criterio di Eisenstein con il primo  $p$ : il coefficiente direttore del nostro polinomio, cioè 1, non è divisibile per  $p$ , ma tutti gli altri coefficienti lo sono, inoltre  $p^2$  non divide il termine noto  $-p$ . Dunque le ipotesi del criterio sono soddisfatte e  $x^n - p$  è irriducibile. Vediamo così che per ogni intero positivo  $n$  esistono in  $\mathbb{Q}[x]$  polinomi irriducibili di grado  $n$ .

Altri esempi di polinomi la cui irriducibilità segue dal criterio di Eisenstein sono  $3x^{10} - 15x^7 + 20x^5 + 5x^2 - 10$  (si può applicare il criterio ponendo  $p = 5$ ) e  $7x^4 + 6x^3 + 12x - 30$  (si può applicare il criterio ponendo  $p = 2$  o anche ponendo  $p = 3$ ). Naturalmente il fatto che non si possa applicare il criterio di Eisenstein ad un polinomio  $f$  non comporta affatto che  $f$  sia riducibile. Ad esempio, al polinomio  $x^3 + 2x + 1$  non si può applicare il criterio di Eisenstein, perché nessun primo ne divide il termine noto, ma ciononostante questo polinomio è irriducibile in  $\mathbb{Q}[x]$  (vedi più avanti l'[Esempio 32](#), per una giustificazione di questo fatto).

Torniamo ora al primo dei due punti considerati all'inizio di questa sezione: in che modo possiamo cercare divisori di un polinomio? Il metodo più semplice, quando è applicabile, è quello fornito dal teorema di Ruffini. Se di un polinomio  $f$  conosciamo una radice  $c$  allora  $f$  è divisibile per  $x - c$ . Dividendo  $f$  per  $x - c$  otteniamo un polinomio  $f_1$  tale che  $f = (x - c)f_1$ . Se stiamo ricercando una fattorizzazione in irriducibili di  $f$  basterà allora trovare una fattorizzazione in irriducibili di  $f_1$  ed aggiungere a questa il fattore  $x - c$ . Ad esempio,  $f = x^3 - 1 \in \mathbb{Q}[x]$  ha chiaramente 1 come radice; possiamo allora dividere  $f$  per  $x - 1$  ottenendo il quoziente  $x^2 + x + 1$ , allora  $f = (x - 1)(x^2 + x + 1)$ . Poiché  $x^2 + x + 1$  non ha radici in  $\mathbb{Q}$  (non ne ha neanche in  $\mathbb{R}$ ) ed ha grado due,  $x^2 + x + 1$  è irriducibile in  $\mathbb{Q}[x]$  per la [Proposizione 23](#); dunque la fattorizzazione ottenuta è la fattorizzazione in irriducibili monici di  $f$  in  $\mathbb{Q}[x]$ .

La ricerca di radici (in  $\mathbb{Q}$ ) di polinomi in  $\mathbb{Q}[x]$  è semplificata enormemente da questo semplice risultato:

**Proposizione 29.** *Sia  $f = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \in \mathbb{Z}[x]$ , con  $a_n \neq 0$ . Allora ogni radice di  $f$  in  $\mathbb{Q}$  si scrive come frazione  $u/v$ , dove  $u$  e  $v$  sono interi coprimi,  $u$  divide  $a_0$  e  $v$  divide  $a_n$ .*

*Dimostrazione.* Ogni numero razionale si può scrivere come frazione ridotta, quindi nella forma  $u/v$ , dove  $u$  e  $v$  sono interi coprimi ( $v \neq 0$ ). Se una tale frazione  $u/v$  è radice di  $f$  allora  $0 = f(u/v) = \sum_{i=0}^n a_i(u/v)^i$ . Moltiplicando per  $v^n$  otteniamo:

$$a_0v^n + a_1uv^{n-1} + a_2u^2v^{n-2} + \cdots + a_{n-2}u^{n-2}v^2 + a_{n-1}u^{n-1}v + a_nu^n = 0.$$

Ora, escluso (per il momento) il primo, tutti gli addendi a primo membro sono multipli di  $u$ . Ma, poiché la loro somma vale 0, il primo addendo  $a_0v^n$  è l'opposto della somma dei rimanenti:  $a_0v^n = -\sum_{i=1}^n a_iu^iv^{n-i}$ , quindi anch'esso è multiplo di  $u$ . Dunque  $u$  divide  $a_0v^n$ . Ma  $u$  è coprimo con  $v$ , quindi con  $v^n$ , dunque  $u$  divide  $a_0$ , come richiesto dall'enunciato. In modo analogo si dimostra che  $v$  divide  $a_n$ : nella somma considerata sopra, escluso l'ultimo addendo  $a_nu^n$  tutti gli altri sono multipli di  $v$ , ma  $a_nu^n$  è l'opposto della somma degli addendi rimanenti, quindi  $v$  divide  $a_nu^n$  e, dal momento che  $v$  e  $u^n$  sono coprimi,  $v$  divide  $a_n$ .  $\square$

Ricordiamo che ogni polinomio in  $\mathbb{Q}[x]$  è associato (in  $\mathbb{Q}[x]$ ) ad un polinomio in  $\mathbb{Z}[x]$ , che avrà le sue stesse radici (in  $\mathbb{Q}$ ). Dunque, volendo ricercare le radici razionali (cioè in  $\mathbb{Q}$ ) di un polinomio  $f \in \mathbb{Q}[x]$  possiamo procedere in questo modo: sostituiamo innanzitutto il polinomio con un suo associato a coefficienti in  $\mathbb{Z}$ , di questo consideriamo il coefficiente direttore  $a_n$  ed il termine noto  $a_0$ ; le radici di  $f$  andranno cercate tra le frazioni ridotte con numeratore divisore di  $a_0$  e denominatore divisore di  $a_n$ . È chiaro che (escluso il caso, banalmente semplificabile, in cui  $a_0 = 0$ ) esiste solo un numero finito di tali frazioni, possiamo allora verificare per ciascuna di esse se è o meno radice del nostro polinomio.

*Esempio 30.* Consideriamo il polinomio  $f = x^4 - 4x^2 + (3/2)x + 3 \in \mathbb{Q}[x]$ . Un suo associato a coefficienti interi è  $2f = 2x^4 - 8x^2 + 3x + 6$ , con coefficiente direttore 2 e termine noto 6. Le frazioni della forma  $u/v$  con  $u$  e  $v$  interi coprimi tali che  $u$  divida 6 e  $v$  divida 2 sono:  $1 = 1/1, 1/2, 2, 3, 3/2, 6$  ed i loro opposti  $-1, -1/2, -2, -3, -3/2, -6$ . Per cercare tutte le radici razionali di  $f$  non dobbiamo fare altro che controllare quali di questi dodici numeri sono radici di  $f$ . Nel nostro caso la verifica diretta mostra che solo  $-2$ , tra questi dodici, è radice. Concludiamo che  $-2$  è l'unica radice di  $f$  in  $\mathbb{Q}[x]$ . Possiamo proseguire lo studio di questo polinomio cercando di fattorizzarlo in prodotto di irriducibili. Usiamo il teorema di Ruffini; dividendo  $f$  per  $x + 2$  (cioè  $x - (-2)$ ) otteniamo  $f = (x + 2)(x^3 - 2x^2 + 3/2)$ . Il secondo fattore  $f_1$  di questo prodotto è associato a  $2f_1 = 2x^3 - 4x^2 + 3$ . Ora, applicando direttamente la [Proposizione 29](#) concluderemmo che le radici di  $f_1$  sono da cercare tra le frazioni ridotte della forma  $u/v$  dove  $u, v \in \mathbb{Z}$ ,  $u$  divide 3 e  $v$  divide 2. In realtà non è necessario esaminare tutte queste frazioni (sono in tutto otto:  $1, 1/2, 3, 3/2$  ed i loro opposti), perché ogni radice di  $f_1$  è anche radice di  $f$  e di tutte queste frazioni, tranne  $-2$ , sappiamo che non sono radici di  $f$ , quindi nemmeno di  $f_1$ . Dobbiamo allora esaminare solo  $-2$ ; si ha  $f_1(-2) = (-2)^3 - 2(-2)^2 + 3/2 \neq 0$ , quindi  $-2$  non è radice di  $f_1$ . Pertanto  $f_1$  non ha radici in  $\mathbb{Q}$ ; poiché  $\nu f_1 = 3$  concludiamo, per la [Proposizione 23](#), che  $f_1$  è irriducibile in  $\mathbb{Q}[x]$ . Dunque una fattorizzazione (l'unica a meno dell'ordine) di  $f$  in prodotto di irriducibili monici in  $\mathbb{Q}[x]$  è  $f = (x + 2)(x^3 - 2x^2 + 3/2)$ .

Una situazione in cui la [Proposizione 29](#) è particolarmente utile è quella in cui il polinomio  $f$  che appare nell'enunciato è monico. In questo caso, infatti, il denominatore  $v$  di una radice  $u/v$  di  $f$  in  $\mathbb{Q}$  deve dividere 1, quindi  $v = 1$  o  $v = -1$ ; ciò comporta che la radice  $u/v$  è un numero intero. Abbiamo allora:

**Corollario 31.** *Sia  $f$  un polinomio monico in  $\mathbb{Z}[x]$ . Allora ogni radice razionale di  $f$  è intera.*

*Esempio 32.* Poco fa abbiamo detto, ma non giustificato, che il polinomio  $f = x^3 + 2x + 1$  è irriducibile in  $\mathbb{Q}[x]$ . Sappiamo che questa affermazione equivale al fatto che  $f$  (che ha grado 3) è privo di radici in  $\mathbb{Q}$ , per la [Proposizione 23](#). In effetti, ogni (eventuale) radice razionale di  $f$  deve essere intera, per il [Corollario 31](#), inoltre, ancora per la [Proposizione 29](#), essa deve dividere il termine noto di  $f$ , che è 1. Dunque gli unici due numeri razionali che potrebbero essere radici di  $f$  sono i divisori interi di 1, cioè 1 e  $-1$ . Ma  $f(1) = 4$  e  $f(-1) = -2$ , quindi nessuno di questi due numeri è radice di  $f$  e così  $f$  è privo di radici. Per questo motivo  $f$  è irriducibile in  $\mathbb{Q}[x]$ .

Un'altra applicazione del [Corollario 31](#) ha a che fare con le radici dei numeri interi. Se  $a \in \mathbb{N}$  e  $n \in \mathbb{N}^*$ , la radice  $n$ -esima di  $a$ ,  $\sqrt[n]{a}$ , è un numero reale la cui  $n$ -esima potenza sia  $a$  (precisamente, l'unico tale numero, se  $n$  è dispari, quello non negativo se  $n$  è pari). Quindi  $\sqrt[n]{a}$  è una radice del polinomio monico  $x^n - a \in \mathbb{Z}[x]$ . Le radici razionali di questo polinomio sono intere, quindi  $\sqrt[n]{a}$  è o intera (ad esempio, se  $n = 2$  e  $a = 4$ ) oppure irrazionale. Questo è un modo per dimostrare che numeri come  $\sqrt{2}$ ,  $\sqrt{3}$  o  $\sqrt[11]{37}$ , che certamente non sono interi, sono irrazionali.

*Esempio 33.* Fattorizziamo in prodotti di invertibili e irriducibili in  $\mathbb{Q}[x]$  i polinomi  $f = 2x^5 - x^3 + 2x^2 - 1$  e  $g = x^5 + x^4 + x^3 + x^2 + x + 1$  dell'[Esempio 9](#). Sappiamo che un loro massimo comun divisore è  $(7/4)(x^3 + 1)$ , quindi  $d = x^3 + 1$  è il loro massimo comun divisore monico. Per fattorizzare  $f$ , conviene iniziare con lo sfruttare questa informazione, che fornisce un divisore non banale, per l'appunto  $d$ , di  $f$ . Dividendo  $f$  per  $d$  abbiamo  $f = df_1$ , dove  $f_1 = 2x^2 - 1$ . Per fattorizzare in irriducibili  $f$  basta dunque fattorizzare separatamente  $d$  e  $f_1$ . Iniziamo con  $d = x^3 + 1$ ; poiché ha grado 3 esso è irriducibile se e solo se non ha radici in  $\mathbb{Q}$ , per la [Proposizione 23](#). La [Proposizione 29](#) (ed il [Corollario 31](#)) ci dicono che le radici razionali di  $d$  sono intere e dividono 1, quindi le sole possibili radici razionali di  $d$  sono 1 e  $-1$ . Ora,  $d(1) = 2$  e  $d(-1) = 0$ , quindi 1 non è radice di  $d$ , ma  $-1$  lo è. Allora, per il teorema di Ruffini,  $d$  è divisibile per  $x - (-1) = x + 1$ . Si ha  $d = (x + 1)(x^2 - x + 1)$ . Le radici razionali di  $h = x^2 - x + 1$  sono radici di  $d$ , la cui unica radice razionale è  $-1$ ; quindi  $-1$  è l'unica possibile radice razionale di  $h$  in  $\mathbb{Q}$ . Ma  $-1$  non è radice di  $h$ , infatti  $h(-1) = 3$ , quindi  $h$  non ha radici in  $\mathbb{Q}$  ed è così irriducibile per la [Proposizione 23](#). Ovviamente avremmo anche potuto osservare, in alternativa, che  $h$  non ha radici reali, quindi non ha radici razionali, perché il suo discriminante è  $-3 < 0$ . Abbiamo così la fattorizzazione di  $d$  in irriducibili monici:  $d = (x + 1)(x^2 - x + 1)$ . Passiamo ora a  $f_1 = 2x^2 - 1 = 2(x^2 - 1/2)$ . Le radici di  $f_1$  in  $\mathbb{R}$  sono  $1/\sqrt{2}$  e  $-1/\sqrt{2}$ , che sono irrazionali (se  $1/\sqrt{2}$  fosse razionale sarebbe razionale anche il suo reciproco  $\sqrt{2}$ , ma sappiamo che  $\sqrt{2} \notin \mathbb{Q}$ ). Quindi  $f_1$  non ha radici razionali e, essendo di secondo grado, è quindi irriducibile. Mettendo insieme le fattorizzazioni di  $d$  e di  $f_1$  otteniamo così la fattorizzazione di  $f$  come prodotto di un invertibile (il suo coefficiente direttore 2) ed irriducibili monici:  $f = 2(x + 1)(x^2 - x + 1)(x^2 - 1/2)$ . Questa fattorizzazione è unica a meno dell'ordine dei fattori (vedi [Proposizione 19](#)).

Fattorizziamo ora  $g$ . Come per  $f$ , iniziamo col dividere  $g$  per il suo divisore non banale  $d$ , ottenendo  $g = dg_1$ , dove  $g_1 = x^2 + x + 1$ . Abbiamo già la fattorizzazione completa di  $d$ ; non è difficile verificare che  $g_1$  è irriducibile, perché ha secondo grado ed è privo di radici. Quest'ultimo fatto si può verificare o osservando che il discriminante di  $g$  è negativo (quindi  $g$  non ha radici in  $\mathbb{R}$ ), oppure che, per la [Proposizione 29](#), le radici razionali di  $g$  sono da cercare tra 1 e  $-1$ , ma queste non sono radici di  $g$ . La fattorizzazione di  $g$  in prodotto di polinomi irriducibili monici in  $\mathbb{Q}[x]$  è quindi  $g = (x + 1)(x^2 - x + 1)(x^2 + x + 1)$ .

Possiamo anche fattorizzare  $f$  e  $g$  in  $\mathbb{R}[x]$ . Conviene partire dalle fattorizzazioni in invertibili e irriducibili ottenute in  $\mathbb{Q}[x]$ . Nella fattorizzazione  $f = 2(x + 1)(x^2 - x + 1)(x^2 - 1/2)$  il fattore di primo grado  $x + 1$  è ovviamente irriducibile in  $\mathbb{R}[x]$ , i due fattori di secondo grado sono irriducibili in  $\mathbb{R}[x]$  se e solo se sono privi di radici reali. Come già detto,  $h = x^2 - x + 1$  non ha radici reali, quindi  $h$  è irriducibile in  $\mathbb{R}$ , mentre  $x^2 - 1/2$  ha due radici reali,  $1/\sqrt{2} = 2/\sqrt{2}$  e  $-1/\sqrt{2}$ , quindi  $x^2 - 1/2 = (x - 1/\sqrt{2})(x + 1/\sqrt{2})$  per il teorema di Ruffini generalizzato. I due fattori appena trovati hanno grado uno e quindi sono irriducibili in  $\mathbb{R}[x]$ . La fattorizzazione in un invertibile e irriducibili monici di  $f$  in  $\mathbb{R}[x]$  è dunque  $f = 2(x + 1)(x^2 - x + 1)(x - 1/\sqrt{2})(x + 1/\sqrt{2})$ . Invece, entrambi i fattori di secondo grado nella fattorizzazione  $g = (x + 1)(x^2 - x + 1)(x^2 + x + 1)$  di  $g$  sono privi di radici reali, quindi irriducibili anche in  $\mathbb{R}[x]$ , pertanto la stessa fattorizzazione è la fattorizzazione di  $g$  in prodotto di irriducibili monici in  $\mathbb{R}[x]$ .

A proposito dell'uso del teorema di Ruffini per ottenere fattorizzazioni in  $\mathbb{Q}[x]$ ,  $\mathbb{R}[x]$  o  $\mathbb{C}[x]$ , menzionamo per completezza il fatto che, così come esiste una formula che fornisce le radici di un polinomio di secondo grado (in cui appare l'estrazione di una radice quadrata), esistono formule simili, ma un pò più complicate, che forniscono le radici dei polinomi di terzo e quarto grado (in cui appaiono estrazioni di radici terze o quarte, rispettivamente) ma non esistono (meglio: non possono esistere) formule dello stesso tipo che forniscono le radici di polinomi di grado superiore al quarto.

I due esempi conclusivi riguardano polinomi su campi finiti. Soprattutto quando la cardinalità del campo finito  $F$  è piccola il metodo più efficace per la ricerca delle radici di un polinomio  $f \in F[x]$  è spesso la verifica diretta eseguita per ogni elemento, vale a dire il calcolo di  $f(c)$  per ogni elemento  $c$  del campo.

*Esempio 34.* Per alcuni valori del primo  $p$  fattorizziamo  $f_p$ , il polinomio  $f = 2x^5 - x^3 + 2x^2 - 1$  dell'esempio precedente riguardato come polinomio a coefficienti in  $\mathbb{Z}_p$ . Tra le applicazioni della proprietà universale viste nella [Sezione 2](#) ricordiamo l'omomorfismo suriettivo  $\bar{\epsilon}_p$  che ad ogni polinomio in  $\mathbb{Z}[x]$  associa il polinomio stesso riguardato come polinomio a coefficienti in  $\mathbb{Z}_p[x]$ . Il fatto che  $\bar{\epsilon}_p$  sia un omomorfismo permette di ‘tradurre’ fattorizzazioni di un polinomio in  $\mathbb{Z}[x]$  in fattorizzazioni della sua immagine in  $\mathbb{Z}_p[x]$ : per il nostro  $f$ , se  $g, h \in \mathbb{Z}[x]$  sono tali che  $f = gh$ , allora  $f^{\bar{\epsilon}_p} = g^{\bar{\epsilon}_p}h^{\bar{\epsilon}_p}$ .

Scriviamo allora  $f$  come prodotto di polinomi a coefficienti interi e irriducibili in  $\mathbb{Z}[x]$ , utilizzando quanto ottenuto nell'esempio precedente:  $f = (x+1)(x^2-x+1)(2x^2-1)$ . Per ogni primo  $p$  abbiamo  $f_p = (x+1)^{\bar{\varepsilon}_p}(x^2-x+1)^{\bar{\varepsilon}_p}(2x^2-1)^{\bar{\varepsilon}_p}$ . Per ottenere una fattorizzazione in prodotto di polinomi irriducibili in  $\mathbb{Z}_p[x]$  si devono allora ulteriormente fattorizzare in prodotto di irriducibili (in  $\mathbb{Z}_p[x]$ ) i tre fattori  $h_{p,1} = (x+1)^{\bar{\varepsilon}_p}$ ,  $h_{p,2} = (x^2-x+1)^{\bar{\varepsilon}_p}$  e  $h_{p,3} = (2x^2-1)^{\bar{\varepsilon}_p}$ . Non c'è nessun problema per il primo fattore, che è di grado 1 e quindi irriducibile qualsiasi sia il primo  $p$ , vanno invece considerati con maggiore attenzione gli altri due fattori, che vanno studiati considerando separatamente i valori di  $p$  a cui siamo interessati. Qui consideriamo i primi minori o uguali a 7:

- $p = 2$ :  $f_2 = x^3 + \bar{1} \in \mathbb{Z}_2[x]$  ha grado 3. Il terzo dei fattori appena presi in esame, infatti, in questo caso si riduce a  $[1]_2$ :  $h_{2,3} = (\bar{2}x^2 - \bar{1})^{\bar{\varepsilon}_p} = \bar{1}$ . Il secondo fattore  $h_{2,2} = x^2 + x + \bar{1}$  è privo di radici in  $\mathbb{Z}_2[x]$ , infatti  $h_{2,2}([0]_2) = h_{2,2}([1]_2) = [1]_2 \neq [0]_2$ . Quindi, per la [Proposizione 23](#),  $h_{2,2}$  è irriducibile in  $\mathbb{Z}_2[x]$  (vedi anche l'esempio successivo). La fattorizzazione di  $f_2$  in prodotto di irriducibili in  $\mathbb{Z}_2[x]$  è dunque  $f_2 = (x+\bar{1})(x^2+x+\bar{1})$ .
- $p = 3$ : Se  $p > 2$ , quindi anche nel caso  $p = 3$  che consideriamo ora,  $\nu f_p = 5$ . Sia  $h_{3,2}$  che  $h_{3,3}$  hanno grado 2, ricerchiamone le (eventuali) radici in  $\mathbb{Z}_3$ . Gli elementi di  $\mathbb{Z}_3$  sono  $[0]_3$ ,  $[1]_3$ , e  $[-1]_3$ , abbiamo  $h_{3,2}([0]_3) = h_{3,2}([1]_3) = [1]_3 \neq [0]_3 = h_{3,2}([-1]_3)$ , quindi  $[-1]_3$  è l'unica radice di  $h_{3,2}$  in  $\mathbb{Z}_3$ . Dal momento che  $\nu h_{3,2} = 2$ , allora  $h_{3,2}$  è riducibile (è divisibile per  $x+\bar{1}$ , per il teorema di Ruffini) ed è il prodotto di due polinomi di grado 1, che possiamo anche scegliere monici perché  $h_{3,2}$  è monico, dunque  $h_{3,2} = (x+\bar{1})(x-c)$  dove  $c$  è una radice di  $h_{3,2}$ . Ma  $[-1]_3$  è l'unica radice di  $h_{3,2}$  in  $\mathbb{Z}_3$  quindi  $c = [-1]_3$  ed allora  $h_{3,2} = (x+\bar{1})^2 \in \mathbb{Z}_3[x]$  (cosa che, ovviamente si può anche verificare direttamente: in  $\mathbb{Z}_3[x]$  si ha  $(x+\bar{1})^2 = x^2 + \bar{2}x + \bar{1} = x^2 - x + \bar{1} = h_{3,2}$ ). Consideriamo ora  $h_{3,3} = -(x^2 + \bar{1})$ ; questo non ha radici in  $\mathbb{Z}_3$ , infatti  $h_{3,3}([0]_3) = [-1]_3$  e  $h_{3,3}([1]_3) = h_{3,3}([-1]_3) = [1]_3$ . Dunque  $h_{3,3}$  è irriducibile e la fattorizzazione di  $f_3$  nel prodotto di un invertibile ed irriducibili monici è  $f_3 = (-\bar{1})(x+\bar{1})^3(x^2 + \bar{1})$ .
- $p = 5$ : Calcolando  $h_{5,2}(c) = c^2 - c + [1]_5$  per ogni  $c \in \mathbb{Z}_5$  verifichiamo rapidamente che  $h_{5,2}([0]_5) = h_{2,5}([1]_5) = [1]_5$ ,  $h_{5,2}([-1]_5) = [3]_5 = [-2]_5 = h_{5,2}([2]_5)$  e  $h_{5,2}([-2]_5) = [2]_5$ . Quindi  $h_{5,2}$  non ha radici in  $\mathbb{Z}_5$  e la [Proposizione 23](#) ne garantisce l'irriducibilità. Per quanto riguarda  $h_{5,3}$  abbiamo poi  $h_{5,3} = \bar{2}x^2 - \bar{1} = \bar{2}x^2 + \bar{4} = \bar{2}(x^2 + \bar{2}) = \bar{2}(x^2 - \bar{3})$ . Come si verifica subito,  $[3]_5$  non è un quadrato in  $\mathbb{Z}_5$  (infatti  $[0]_5^2 = [0]_5$ ,  $[1]_5^2 = [-1]_5^2 = [1]_5$  e  $[2]_5^2 = [-2]_5^2 = [4]_5$ ), quindi anche  $h_{5,3}$  è privo di radici in  $\mathbb{Z}_5$  ed è così irriducibile in  $\mathbb{Z}_5[x]$ . Dunque, la fattorizzazione di  $f_5$  nel prodotto di un invertibile ed irriducibili monici è  $f_5 = \bar{2}(x+\bar{1})(x^2 - x + \bar{1})(x^2 + \bar{2})$ .
- $p = 7$ : Ragionando come nei casi precedenti, cerchiamo le radici di  $h_{7,2}$ . Scopriamo che  $[-2]_7$  e  $[3]_7$  sono radici di  $h_{7,2}$ . Abbiamo poi  $h_{7,3} = \bar{2}x^2 - \bar{1} = \bar{2}x^2 + \bar{6} = \bar{2}(x^2 + \bar{3}) = \bar{2}(x^2 - \bar{4}) = \bar{2}(x+\bar{2})(x-\bar{2})$  (quindi  $h_{7,3}$  ha radici  $[2]_7$  e  $[-2]_7$ ). Allora tutti i fattori irriducibili di  $f_7$  hanno grado 1; la fattorizzazione nel prodotto di un invertibile ed irriducibili monici è  $f_7 = \bar{2}(x+\bar{1})(x-\bar{3})(x+\bar{2})^2(x-\bar{2})$ .

*Esempio 35.* Possiamo usare i risultati di queste due ultime sezioni per elencare, uno per uno, tutti i polinomi irriducibili di assegnato grado in  $\mathbb{Z}_2[x]$ . Per qualsiasi campo  $K$  e per ogni  $n \in \mathbb{N}$  i polinomi di grado  $n$  in  $K[x]$  sono tutti quelli della forma  $a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$  dove  $a_0, a_1, \dots, a_{n-1} \in K$  e  $a_n \in K \setminus \{0_K\}$ . Nel nostro caso, in cui  $K = \mathbb{Z}_2$ , richiedere  $a_n \in \mathbb{Z}_2 \setminus \{0_{\mathbb{Z}_2}\}$  significa richiedere  $a_n = [1]_2$ , dunque tutti i polinomi non nulli in  $\mathbb{Z}_2[x]$  sono monici. Abbiamo allora:

- esattamente due polinomi di grado uno:  $x = x + \bar{0}$  e  $x + \bar{1}$ . Essendo di grado uno, questi sono irriducibili.
- I polinomi di grado due sono quelli della forma  $x^2 + a_1 x + a_0$ , dove  $a_1$  e  $a_0$  possono essere  $\bar{0}$  o  $\bar{1}$ . Abbiamo dunque quattro polinomi di grado due:  $x^2$ ,  $x^2 + x$ ,  $x^2 + \bar{1}$ ,  $x^2 + x + \bar{1}$ . Tra questi sono irriducibili tutti e soli quelli privi di radici. I primi due hanno  $\bar{0}$  come radice, il terzo ha  $\bar{1}$  come radice, il quarto non ha né  $\bar{0}$  né  $\bar{1}$  come radice, quindi è privo di radici ed è così irriducibile in  $\mathbb{Z}_2[x]$ , l'unico irriducibile di grado 2.
- Passiamo ai polinomi di grado tre: questi hanno la forma  $x^3 + a_2 x^2 + a_1 x + a_0$ , dove  $a_2, a_1$  e  $a_0$  possono essere scelti tra  $\bar{0}$  e  $\bar{1}$ . Abbiamo così otto polinomi di grado tre; tra questi quelli privi di radici in  $\mathbb{Z}_2$ , cioè irriducibili in  $\mathbb{Z}_2[x]$ , sono:  $x^3 + x^2 + \bar{1}$ ,  $x^3 + x + \bar{1}$  e nessun altro.
- Per i polinomi di grado due o tre abbiamo usato la [Proposizione 23](#); questa non può essere più utilizzata nel caso dei polinomi di quarto grado. Dei sedici polinomi di quarto grado esattamente quattro sono privi di radici in  $\mathbb{Z}_2$ , essi sono:  $x^4 + x^3 + x^2 + x + \bar{1}$ ,  $x^4 + x^3 + \bar{1}$ ,  $x^4 + x^2 + \bar{1}$  e  $x^4 + x + \bar{1}$ . Un polinomio  $f$  di quarto grado (su un campo qualsiasi) che sia riducibile ma non abbia radici deve avere una fattorizzazione non banale del tipo  $f = gh$  in cui  $4 = \nu g + \nu h$  ma  $\nu g \neq 1 \neq \nu h$ , perché se  $f$  avesse un fattore di grado 1 allora avrebbe una radice ([Proposizione 21](#)), quindi deve aversi  $\nu g = \nu h = 2$ . Inoltre, poiché  $f$  è privo di radici anche  $g$  ed  $h$  sono privi di radici, quindi irriducibili. Dunque, un polinomio di quarto grado a coefficienti in un campo è irriducibile se e solo se è privo di radici e non è il prodotto di due polinomi irriducibili di grado due. Nel caso del campo  $\mathbb{Z}_2$ , che stiamo considerando, abbiamo visto che esiste solo un polinomio irriducibile di grado due:  $x^2 + x + \bar{1}$ . Allora i polinomi irriducibili di grado quattro in  $\mathbb{Z}_2[x]$  sono tutti e soli quelli privi di radici ad eccezione di  $(x^2 + x + \bar{1})^2$ . Poiché, come si vede rapidamente,  $(x^2 + x + \bar{1})^2 = x^4 + x^2 + \bar{1}$ , concludiamo che i polinomi irriducibili di grado quattro in  $\mathbb{Z}_2[x]$  sono:  $x^4 + x^3 + x^2 + x + \bar{1}$ ,  $x^4 + x^3 + \bar{1}$  e  $x^4 + x + \bar{1}$ .

Abbiamo così stabilito che in  $\mathbb{Z}_2[x]$  esistono esattamente due polinomi irriducibili di grado 1, uno di grado 2, due di grado 3, tre di grado 4. Si potrebbe continuare, con lo stesso metodo, ad elencare i polinomi irriducibili in  $\mathbb{Z}_2[x]$  di gradi maggiori. Ad esempio, osservando che i polinomi irriducibili di grado cinque a coefficienti in un campo

sono quelli privi di radici che non siano prodotto di un polinomio di grado due ed uno di grado tre si può arrivare a concludere che i polinomi irriducibili di grado cinque in  $\mathbb{Z}_2[x]$  sono esattamente sei.

Si può anche ripetere l'esercizio per altri campi finiti. In questo caso non è più vero che i polinomi non nulli sono tutti monici, ma per trovare tutti quelli irriducibili basta comunque trovare gli irriducibili monici ed aggiungere poi alla lista i loro associati. Ad esempio, i polinomi irriducibili di secondo grado in  $\mathbb{Z}_3[x]$  sono  $x^2 + \bar{1}$ ,  $x^2 + x - \bar{1}$ ,  $x^2 - x - \bar{1}$  (che sono i polinomi monici di grado due privi di radici) ed i loro opposti, che sono i loro altri associati.

# Polinomi

Sia  $A$  un anello commutativo unitario.

anello di polinomi:  $A[x]$  un anello commutativo unitario tale che:

1)  $A$  sia un sottoanello unitario di  $A[x]$ .

2)  $x \in A[x]$ .

3)  $\forall f \in A[x] \quad (\exists! (a_i)_{i \in \mathbb{N}} \in A^{\mathbb{N}} \quad (\exists n \in \mathbb{N} \quad (f = a_0 + a_1x + a_2x^2 + \dots + a_nx^n \quad \& \quad \forall i \in \mathbb{N} \quad (i > n \rightarrow a_i = 0_A)))$ )

Esempio:

$$A = \mathbb{Z}$$

$$x = \sqrt{2}$$

$$\underline{a} = (a_i)_{i \in \mathbb{N}} \quad a_0 = 0 = a_1 \quad \forall i \in \mathbb{N} \setminus \{2\} \quad \& \quad a_2 = 1$$
$$\underline{b} = (b_i)_{i \in \mathbb{N}} \quad b_0 = 2 \quad \forall i \in \mathbb{N}^* \quad (b_i = 0)$$
$$\rightarrow 0 + 0(\sqrt{2}) + 1(\sqrt{2})^2 \dots = 2$$
$$\rightarrow 2 + 0(\sqrt{2}) + 0(\sqrt{2})^2 \dots = 2$$

Mancava l'unicità!

Si verifica, invece, per  $\pi$  e  $e$ .

$$f = \sum_{i=0}^n a_i x^i \Rightarrow \text{se } n \geq m \quad \forall i \in \{0, 1, \dots, m\} \quad a_i = b_i$$
$$f = \sum_{i=0}^m b_i x^i \quad \forall i \in \{m+1, \dots, n\} \quad a_i = 0_A$$

Ogni elemento di un polinomio è polinomio di costanti.

$0_A = 0_{A[x]}$  polinomio nullo : successione di coefficienti:  $(0_i)_{i \in \mathbb{N}}$

Se  $f \in A[x] - \{0_A\}$  e  $(a_i)_{i \in \mathbb{N}}$  è la successione dei coefficienti di  $f$

$S_f = \{i \in \mathbb{N} \mid a_i \neq 0_A\}$  è finito e non vuoto, quindi ha massimo.

$\max S_f$  si chiama GRADO di  $f$ :  $\circ f$  ( $\circ S_f$ ,  $\deg f$ )

Se  $n = \circ f$ ,  $a_n$  è il coefficiente direttore di  $f$ .

Esempio:

$$\text{in } \mathbb{Z}_5[x] \quad f = [2]_5 + [3]_5 x + [10]_5 x^3 \quad \stackrel{\text{coz}}{\circ} f = 1 \quad cd(f) = [3]_5$$

Si pone anche  $cd(0_A) = 0_A$  e  $\circ(0_A) = -\infty$

## Proprietà dei polinomi

- $f$  è monico  $\Leftrightarrow \text{cd}(f) = 1_A$   
In  $\mathbb{Z}_2$ , ogni polinomio non nullo è monico.

## Proprietà universale.

Siamo  $A$  e  $B$  anelli commutativi unitari e  $A[x]$  un anello di polinomi su  $A$  ed una indeterminata  $x$  e sia  $\vartheta$  un omomorfismo di anelli unitari da  $A$  a  $B$

$$\text{Allora: } \vartheta(1_A) = 1_B$$

$$\forall c \in B$$

$$\exists \vartheta_* \text{ omom. di anelli unitari: } A \xrightarrow{\vartheta} B$$
$$A[x] \rightarrow B \text{ tale che: } \vartheta_*|_{A[x]} = \vartheta$$
$$(\vartheta_*)_A = \vartheta \quad \vartheta_*(x) = c$$

$$\vartheta_*(a_i) = \vartheta(a_i) \quad \vartheta_*(x^i) = (\vartheta(x))^i = c^i$$

$$\sum_{i=0}^n \vartheta_*(a_i x^i) = \sum_{i=0}^n \vartheta_*(a_i) \vartheta_*(x^i)$$
$$\vartheta_*: \sum_{i=0}^n a_i x^i \in A[x] \mapsto \sum_{i=0}^n \vartheta(a_i) c^i \in B$$

## Applicazioni

1) Siano  $A[x]$  e  $A[y]$  anelli di polinomi.

Poniamo  $B = A[y]$ ; come  $\vartheta$  scelgo  $A \hookrightarrow A[y]$ ; come  $c$  scelgo  $y$ .

$$\begin{array}{ccc} A & \xhookrightarrow{\quad} & A[y] \\ \downarrow & & \nearrow \vartheta_x : \sum_{i=0}^n a_i x^i \mapsto \sum_{i=0}^n a_i y^i \\ A[x] & & \text{omomorfismo di anelli unitari} \end{array}$$

Possiamo fare l'inverso:

$$\begin{array}{ccc} A & \xhookrightarrow{\quad} & A[x] \\ \downarrow & & \nearrow \Psi_x : \sum_{i=0}^n a_i y^i \mapsto \sum_{i=0}^n a_i x^i \\ A[y] & & \text{omomorfismo anche questo!} \end{array}$$

Allora  $\Psi_x = \vartheta_x^{-1}$  quindi  $\vartheta_x$  è un isomorfismo.

2) Poniamo  $B = A$  e  $\vartheta = \text{id}_A$

$\forall c \in A$

$$\begin{array}{ccc} A & \xrightarrow{\text{id}_A} & A \\ \downarrow & & \nearrow f = \sum_{i=0}^n a_i x^i \mapsto \sum_{i=0}^n a_i c^i = f(c) \\ A[x] & & \text{omomorfismo di sostituzione} \end{array} \quad \begin{array}{l} c \mapsto f(c) \\ c \mapsto g(c) \end{array}$$

$$\text{Se } c = 0_A \quad \sum_{i=0}^n a_i x^i \mapsto \underbrace{\sum_{i=0}^n a_i 0_A^i}_{= 0_A}$$

$$3) \text{ Sia } m \in \mathbb{N}^* \quad A = \mathbb{Z} \quad \varepsilon_m : \mathbb{Z} \xrightarrow{\text{proiezione canonica}} \mathbb{Z}_m \quad \mathbb{Z}_m \xrightarrow{\vartheta} \mathbb{Z}_m[x]$$

$$\begin{array}{ccc} \mathbb{Z} & \xrightarrow{\vartheta} & \mathbb{Z}_m[x] \\ \downarrow & & \nearrow f = \sum_{i=0}^n a_i x^i \mapsto \sum_{i=0}^n \bar{a}_i x^i \\ \mathbb{Z}[x] & & \text{omomorfismo suriettivo} \end{array} \quad \begin{array}{c} \vartheta \circ \varepsilon_m = \vartheta \cdot \mathbb{Z} \xrightarrow{\vartheta \circ \varepsilon_m} \mathbb{Z}_m[x] \\ a \mapsto \bar{a} \end{array}$$

$$f = 2 + 3x + 10x^3 \in \mathbb{Z}[x] \quad f_{\bar{a}} = \bar{2} + \bar{3}x + \bar{10}x^3 \in \mathbb{Z}_5[x] = \bar{2} + \bar{3}x$$

Non conserva il grado

Suriettivo ma non iniettivo.

## Operazioni e gradi

$$\forall f, g \in A[x] - \{O_A\}$$

poniamo  $n = \nu f$ ,  $m = \nu g$   
 $a = \text{cd}(f)$ ,  $b = \text{cd}(g)$

$$f = \sum_{i=0}^n a_i x^i \quad a = a_n \neq O_A$$

$$g = \sum_{i=0}^m b_i x^i \quad b = b_m \neq O_B$$

$$\nu(f+g) \leq \max\{n, m\} = M$$

Precisamente:

$$(a_i)_{i \in \mathbb{N}} \text{ succ. coeff. of } f$$

$$(b_i)_{i \in \mathbb{N}} \text{ succ. coeff. of } g$$

$$f+g = \sum_{i=0}^{\max\{n, m\}} (a_i + b_i) x^i$$

$$n \neq m \Rightarrow \nu(f+g) = M$$

$$n = m \wedge a+b \neq O_A \Downarrow$$

$$n = m \wedge a+b = O_A \Rightarrow \nu(f+g) < M$$

$$\nu(f-g) \leq M$$

$$\nu(f-g) < M \Leftrightarrow n = m \wedge a = b$$

$$fg = \sum_{i=0}^{n+m} c_i x^i \quad \text{dove } \forall i \in \mathbb{N} \quad c_i = \sum_{j=0}^i a_j b_{i-j}$$

$$a_0 b_0 + (a_0 b_1 + a_1 b_0)x + (a_0 b_2 + a_1 b_1 + a_2 b_0)x^2 + \dots + a_n b_m x^{n+m}$$

$$\nu(fg) \leq n+m$$

$$\nu(fg) = n+m \Leftrightarrow ab \neq O_A$$

# Algebra

## Lezione 12/12



## Numeri primi

$\forall n \in \mathbb{N}, n > 1 \Rightarrow \exists p \text{ primo positivo:}$   
 $n \text{ non primo} \quad p | n \wedge p \leq \sqrt{n}$

DIM Se  $n$  non è primo, esiste un  $p$  che divide  $n$ . Sia  $p$  il minimo tra questi primi. Allora:  
 $\exists t \in \mathbb{N}^* (n = pt)$ . ovviamente  $t > 1$ . Allora esiste un primo  $q$  divisor di  $t$ . Quindi:  $q | t$ .  
Per la scelta di  $p$ ,  $p \leq q$ . Ma  $q \leq t \Rightarrow p \leq t$ . Quindi:  $p^2 \leq pt = n$ , cioè  $p \leq \sqrt{n}$ .

## Polinomi

$A$  anello comm. unitario.  $A[x]$  anello di polinomi a coefficienti in  $A$  nell'indeterminata  $x$ .

Sia  $f \in A[x]$

① se  $cd(f)$  è cancellabile in  $A$ , allora:

1.  $\forall g \in A[x]$  vale R.A.G. per  $f \cdot g$ .
2.  $f$  è cancellabile in  $A[x]$ .

N.B. R.A.G. = regola di addizione dei gradi

② Sono equivalenti:

1.  $A$  è un dominio di integrità.

2.  $\forall f, g \in A[x] \quad (\deg(fg) = \deg f + \deg g)$ .

3.  $A[x]$  è un dominio di integrità.

$$\Rightarrow \mathcal{U}(A[x]) = \mathcal{U}(A)$$

$$1 \Rightarrow 2 \quad \forall f \in A[x]$$

$f \neq 0_A \Rightarrow cd f$  è cancellabile

$f = 0_A \Rightarrow$  vale RAG per  $f \cdot g$ , qualsiasi sia  $g$

Esempio:  $\mathcal{U}(\mathbb{Z}[x]) = \mathcal{U}(\mathbb{Z}) = \{1, -1\}$

$\forall$  campo  $K \quad \mathcal{U}(K[x]) = K - \{0_K\}$

DIM Ovviamente  $\mathcal{U}(A) \subseteq \mathcal{U}(A[x])$

$\forall f \in \mathcal{U}(A)$  se  $g$  è l'inverso di  $f$  in  $A$ ,  $fg = 1_A = 1_{A[x]}$  quindi  $g$  è l'inverso di  $f$  in  $A[x]$ .

Viceversa, sia  $f \in \mathcal{U}(A[x])$  e  $g$  il suo inverso in  $A[x]$ . Allora  $fg = 1_A$ . Per la RAG  $v_f + v_g = v(1_A) = 0$ .

Dunque  $v_f = v_g = 0$ , quindi  $f, g \in A$ .

Allora  $f \in A$  e  $g$  è l'inverso di  $f$  in  $A$ , quindi  $f \in \mathcal{U}(A)$ .

③ Se  $cd f$  è cancellabile e  $f \in \mathcal{U}(A[x])$  allora  $f \in \mathcal{U}(A)$ .

Se  $|A| > 1$ ,  $f \in A[x]$ ,  $v_f > 0$  e  $cd f$  è cancellabile  $\Rightarrow f \notin \mathcal{U}(A[x])$

Esempio:  $x \notin \mathcal{U}(A[x])$

Per ogni scelta di  $A$ ,  $A[x]$  non è un campo.

## Teorema

Siano  $f, g \in A[x]$ . Se  $b = cd(g) \in U(A)$ :

$$\exists! (q, r) \in A[x] \times A[x] \quad (f = qg + r \wedge \deg r < \deg g)$$

Se  $A$  è un campo, la condizione  $cd(g) \in U(A)$  equivale a:  $g \neq O_A$

### DIM

#### ① Unicità

Siano  $q_1, q_2, r_1, r_2 \in A[x]$

$$f = gq_1 + r_1 = gq_2 + r_2 \quad \deg r_1 < \deg g \quad \text{e} \quad \deg r_2 < \deg g.$$

Visto RAGIONAMENTO  
 $cd(g) \in U(A) \rightarrow g(q_1 - q_2) = r_2 - r_1$

$$\text{Ra grado } v_g + v(q_1 - q_2) \quad \text{Ra grado} < v_g$$

$$v_g + v(q_1 - q_2) < v_g$$

$$\text{Allora } v(q_1 - q_2) = -\infty, \text{ cioè } q_1 - q_2 = O_A \Rightarrow q_1 = q_2.$$

$$\text{e } r_2 - r_1 = (q_1 - q_2)g = O_A \cdot g = O_A \Rightarrow r_2 = r_1$$

#### ② Esistenza

poniamo  $a = cd(f)$ ,  $b = cd(g)$ ,  $n = vf$ ,  $m = vg$

Caso I  $n < m$ . Basta porre  $q = O_A$  e  $r = f$

Caso II  $n \geq m$ .  $f = ab^{-1}x^{n-m}g = f_1 \quad vf_1 < m$   
Ra  $cd = a$  e grado = n

Esempio:  $A = \mathbb{Q}$   $f = 3x^4 + 2$   $g = 2x - 1$   $\begin{matrix} n=4 & a=3 \\ m=1 & b=2 \end{matrix}$

$$\begin{array}{r} 3x^4 \\ -3x^4 + \frac{3}{2}x^3 \\ \hline \frac{3}{2}x^3 + 2 \end{array} \quad \begin{array}{r} 2x-1 \\ \hline \frac{3}{2}x^3 + 2 \end{array}$$
  
$$f_1 \rightarrow \begin{array}{r} 3x^4 \\ -\frac{3}{2}x^3 + 2 \\ \hline \frac{3}{2}x^3 + 2 \end{array}$$
  
$$f_2 \rightarrow \begin{array}{r} 3x^4 + 2 \\ \hline \frac{3}{2}x^3 + 2 \end{array}$$

$$f_1 = f - \frac{3}{2}x^3 \cdot g = f - (3x^4 - \frac{3}{2}x^3) = 2 + \frac{3}{2}x^3$$

Procedendo in questo modo:

$$f_1 = g \cdot \bar{q} + \bar{r} \quad f = ab^{-1}x^{n-m}g + f_1$$

$$f = ab^{-1}x^{n-m}g + g \cdot \bar{q} + \bar{r} = (\underbrace{ab^{-1}x^{n-m} + \bar{q}}_{=q})g + \bar{r} = r$$

## Applicazioni polinomiali

$\forall f \in A[x] \quad \forall c \in A$   
posto  $f = \sum_{i=0}^n a_i x^i \quad f(c) = \sum_{i=0}^n a_i c^i \quad \tilde{f}: c \in A \mapsto f(c) \in A$   
 $(\forall i \in \{0, 1, \dots, n\} \quad (a_i \in A))$  applicazione polinomiale definita da  $f$

$c$  è radice di  $f: \Leftrightarrow f(c) = 0_A$

$\forall f, g \in A[x]$   
 $f | g \quad \Rightarrow$  ogni radice di  $f$  è radice di  $g$   
 $f \sim g \quad \Rightarrow$   $f$  e  $g$  hanno esattamente le stesse radici associate.

Se  $A$  è un dominio di integrità

$\forall c \in A$   
 $c$  è radice di  $f \quad \Leftrightarrow \quad$  ovvero  
 $c$  è radice di  $f g \quad \Leftrightarrow \quad 0_A = (fg)(c) \Rightarrow f(c) \cdot g(c) = 0_A \quad \downarrow \quad$   $A$  integro  
 $\Rightarrow f(c) = 0_A \quad \vee \quad g(c) = 0_A$

## Teorema del resto

$\forall f \in A[x] \quad \forall c \in A$  nella divisione di  $f$  per  $x-c$  è  $f(c)$

DIM  $f = q \cdot (x-c) + r$   $\forall r \in V(x-c) = 1$ , allora  
omomorf.  
sostituz.  $f(c) = q(c) \cdot (x-c)(c) + r(c) = r(c) = r$   
 $(x-c)(c) = c - c = 0_A$   $\uparrow$  perché  $r \in A$

## Teorema di Ruffini

$\forall f \in A[x] \quad \forall c \in A \quad c$  è radice di  $f \Leftrightarrow x-c | f$

## Teorema di Ruffini generalizzato

Siano  $A$  un dominio di integrità unitario,

$n \in \mathbb{N}^*$   $c_1, c_2, \dots, c_n$  elementi di  $A$  a due a due distinti.  $f \in A_{\text{cxz}}$ .

Sono equivalenti:

$$\textcircled{1} \quad \forall i \in \{1, \dots, n\} \quad f(c_i) = 0_A$$

$$\textcircled{2} \quad \prod_{i=1}^n (x - c_i) \mid f$$

DIM Per induzione

base:  $n=1$   $f$  ok per Ruffini

Sia  $n > 1$  e assumiamo al risultato vero per  $n-1$

valga  $\textcircled{1}$ , allora: per Ruffini è facile  $f(c_n) = 0_A$ ,  $x - c_n \mid f$ .

$\exists h \in A_{\text{cxz}} \quad (f = (x - c_n)h) \quad \forall i \in \{1, \dots, n-1\} \quad f(c_i) = 0_A$  e  $c_i - c_n \neq 0_A$ , quindi  $h(c_i) = 0_A$

per ipotesi di induzione  $\prod_{i=1}^{n-1} (x - c_i) \mid h$   $\exists g \in A_{\text{cxz}} \quad h = g \cdot \prod_{i=1}^{n-1} (x - c_i)$ , che moltiplicato per  $x - c_n$

$$h(x - c_n) = g \cdot \left( \prod_{i=1}^{n-1} (x - c_i) \right) (x - c_n) = g \cdot \prod_{i=1}^n (x - c_i)$$

Algebra

Lezione 14/12



## Teorema di Ruffini generalizzato

Siano  $A$  dominio di integrità unitario,  $f \in A[x]$ ,  $C \subseteq A$ ,  $|C| = n \in \mathbb{N}^*$

Allora:

$$(\forall c \in C \quad (f(c) = 0_A)) \iff \prod_{\substack{c \in C \\ A \subset \mathbb{N}}} (x - c) \mid f$$

### Conseguenze

1) Sia  $0_A \neq f \in A[x]$ . Allora  $f$  ha al massimo  $n$  radici in  $A$ .

Non vale  $\lim x^{2-1} \in \mathbb{Z}_{\geq 2}[x]$ : ha come radici  $i, \bar{i}, \sqrt{2}, -\sqrt{2}$ .

$$A = (P(\mathbb{N}), \Delta, \cap) \quad f = x^2 + x \in A[x] \quad \forall c \in A \quad f(c) = 0_A$$

2) Sia  $A$  un dominio di integrità infinito.

Allora l'applicazione  $f \in A[x] \mapsto \tilde{f} \in \text{Map}(A, A)$  è iniettiva.

Cioè:  $\forall f, g \in A[x] \quad (f = g \iff \tilde{f} = \tilde{g})$

DIM

$$\forall f, g \in A[x]$$

$$\tilde{f} = \tilde{g} \Rightarrow ?$$

$$\text{Posto } h = f - g \quad h(c) = (f - g)(c) = f(c) - g(c) = \tilde{f}(c) - \tilde{g}(c) = 0_A$$

$\forall c \in A \quad c \text{ è radice di } h$ .

Per (1),  $h = 0_A$ , quindi  $f = g$ .

- A dominio di integrità unitario. Allora  $\mathcal{U}(A[x]) = \mathcal{U}(A)$ .  
 $\forall f \in A[x]$  gli associati ad  $f$  in  $A[x]$  sono tutti e soli i polinomi delle forme  $uf$  al variare di  $u$  in  $\mathcal{U}(A)$ , quindi hanno lo stesso grado di  $f$ .

- Se  $A$  è un campo e  $O_A \neq f \in A[x]$   
 $\forall c \in A \setminus \{O_A\}$   $f$  ha esattamente un associato in  $A[x]$  con  $cd = c$ .

DIM Se  $a = cd(f) \in \mathcal{U}(A)$   
 $cd(uf) = ua$  quindi  $uf$  ha coeff. diretto  $c \Leftrightarrow u = a/c$

- Sia  $K$  un campo.

$$\forall f \in K[x] \setminus \{O_K\}$$

polinomi irriconducibili  
 monici

$$f = a \bar{p}_1 \bar{p}_2 \dots \bar{p}_k$$

$$a = cd(f) \in K \setminus \{O_K\}$$

$$f = p_1 p_2 \dots p_k \quad k \in \mathbb{N}^+ \quad \forall i \in \{1, 2, \dots, k\} \quad p_i \text{ è irriducibile.}$$

se  $c_i = cd(p_i)$ ,  $p_i = c_i \bar{p}_i$        $\bar{p}_i$  è monico     $\bar{p}_i \sim p_i$      $\bar{p}_i$  irriducibile.

$$a = \prod_{i=1}^k c_i$$

Sia  $f \in K[x]$       Sia  $n = vf$   
 Allora  $f$  è irriducibile in  $K[x] \Leftrightarrow n > 0$  è:  
 1)  $\rightarrow (\exists g, h \in K[x] \quad f = gh \wedge vg < n \wedge vh < n)$   
 1.5 - equivalente  $\rightarrow (\text{ `` `` `` } \wedge vg > 0 \wedge vh > 0)$

DIM  $f$  irriducibile  $\Rightarrow f \notin \mathcal{U}(K[x]) = K \setminus \{O_K\}$ ,  $f \neq O_K$  dunque  $f \notin K$ , cioè  $n > 0$   
 se  $h, g \in K[x]$  e  $f = gh$ , allora  $g \mid f$ , quindi  $g$  è un dir. banale di  $f$

Quindi valgono 1, 1.5       $vg = 0 \Leftrightarrow$  invertibile ; associato  $\Rightarrow vg = n$

Viceversa, assumiamo  $n > 0$  e valga 1.5

Allora  $f \notin \mathcal{U}(K[x])$ . Se  $g$  è un divisore non banale di  $f$ , allora:

$$\exists h \in K[x] \quad f = gh. \quad Ma \quad g \nmid f, \quad \text{quindi} \quad h \in \mathcal{U}(K[x]) \Rightarrow vh > 0.$$

$$g \notin \mathcal{U}(K[x]) \Rightarrow vg > 0.$$

• A dominio di integrità unitario.

$f \in A[x]$  Se  $f$  ha una radice  $c \in A$ , allora  $x - c$  è un dir. non banale di  $f$  in  $A[x]$ .  
 $\sqrt{f} > 1$

• Sia  $K$  un campo. Sia  $f \in K[x] \setminus \{0_K\}$ , sia  $n = \sqrt{f}$ .

1)  $f$  ha una radice in  $K \iff f$  ha un divisore di grado in  $K[x]$ .

DIM  $h = ax + b \in K[x]$

$a \neq 0_A \Rightarrow -a^{-1}b$  è una radice di  $h$

$n=0 \Rightarrow f$  non ha radici;  $f$  è invertibile, non irriducibile (in  $K[x]$ )

$n=1 \Rightarrow f$  ha una radice;  $f$  è irriducibile

$n > 1 \Rightarrow (f$  irriducibile  $\Rightarrow f$  non ha radici in  $K$ )

$n \in \{2, 3\} \Rightarrow (\text{``} \iff \text{``})$

DIM Sia  $f$  privo di radici; allora  $f$  non ha divisori di grado 1.

Se  $g, h \in K[x] \wedge f = gh$ , allora  $\sqrt{g} + \sqrt{h} = n \Rightarrow \sqrt{g} \neq 1 \neq \sqrt{h}$

Quindi:  $n=2 \Rightarrow \{\sqrt{g}, \sqrt{h}\} = \{0, 2\} \wedge n=3 \Rightarrow \{\sqrt{g}, \sqrt{h}\} = \{0, 3\}$ , quindi  $f$  irriducibile.

Esempio:  $K = \mathbb{Q}$ ,  $(x^2 + 1)^2$  ha grado 4, non ha radici, non è irriducibile.

$K = \mathbb{Z}_2$  polinomi irriducibili in  $\mathbb{Z}_2[x]$

grado 1:  $x, x + \bar{1}$

grado 2:  $x^2 + x + \bar{1}$   $f = x^2 + ax + b$

grado 3:  $x^3 + x^2 + \bar{1}, x^3 + x + \bar{1}$   $f = x^3 + ax^2 + bx + c$

grado 4:  $x^4 + x^3 + x^2 + x + \bar{1}$

$x^4 + x^3 + \bar{1}$ ,  ~~$x^4 + x^2 + \bar{1}$~~ ,  $x^4 + x + \bar{1}$

$\forall f \in \mathbb{Q}[x] (\exists f_1 \in \mathbb{Z}_2[x] (f_1 \sim_{\mathbb{Q}[x]} f))$

## Criterio di Eisenstein

Sia  $f = \sum_{i=0}^n a_i x^i$  un polinomio di grado  $n$ , in  $\mathbb{Z}[x]$ .

Se esiste un numero primo  $p$  tale che

- 1)  $p \nmid a_0$
- 2)  $\forall i \in \{0, 1, \dots, n-1\} \quad p \mid a_i$
- 3)  $p \nmid a_n$

$\forall n \in \mathbb{N}^*$   $x^n - p$  è  
irriducibile in  $\mathbb{Q}(x)$ .

Allora  $f$  è irriducibile in  $\mathbb{Q}(x)$ .

Valga il criterio, sia  $c$  una radice di  $f$  in  $\mathbb{Q}$ .

Scritto  $c$  come frazione ridotta  $\frac{u}{v}$  (cioè:  $u, v$  sono interi, coprimi e  $v \neq 0$ ), si ha:

$$v \mid a_n \quad \wedge \quad u \mid a_0$$

DIM  $0 = f(c) = a_n \frac{u^n}{v^n} + a_{n-1} \frac{u^{n-1}}{v^{n-1}} \dots + a_1 \frac{u}{v} + a_0$  multipli di  $v$

$$0 = f(c) = \underbrace{a_n u^n + a_{n-1} u^{n-1} v + \dots + a_1 u v^{n-1}}_{\text{multipli di } u} + a_0 v^n$$

$$u \mid a_0 v^n \quad u \text{ e } v^n \text{ coprimi} \Rightarrow u \mid a_0$$

$$v \mid a_n u^n \quad v \text{ e } u^n \text{ coprimi} \Rightarrow v \mid a_n$$

Esempio:

$$2x^4 + \dots + 3 \quad v|2 \Rightarrow v \in \{1, 2\}$$

$$v \nmid 2 \quad u|3 \Rightarrow u \in \{\pm 1, \pm 3\}$$

$u, v$  coprimi.

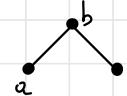
# Grafi

Grafo semplice:

Coppia ordinata  $(V, L)$  dove  $V$  è un insieme (non vuoto)  
 $\& L \subseteq P_2(V)$

elementi di  $V$ : vertici (o nodi)

elementi di  $L$ : lati (o archi)



$$\left( \underbrace{\{a, b, c\}}_V, \underbrace{\{\{a, b\}, \{b, c\}\}}_L \right)$$

Multigrafo semplice:

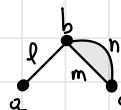
$(V, L, f)$

$V \neq \emptyset$

$f: L \rightarrow P_2(V)$  tale che:

$$l \mapsto \{a, b\}$$

$$n \mapsto \{b, c\} \quad m \mapsto \{b, c\}$$



# Algebra

## Lezione 16/12



# Grafi

Grafo semplice coppia ordinata  $(V, L)$  dove  $V$  è un insieme e  $L \subseteq P_2(V)$

oppure  $(V, p)$

$$\forall V \quad A = \{ p \in \text{Rel}(V) \mid p \text{ è antisimmetrica} \}$$

biettiva:  $L \in P(P_2(V)) \mapsto p_L \in A \quad (\forall a, b \in V \quad a p_L b \Leftrightarrow \{a, b\} \in L)$

con inversa:  $p \in A \mapsto \{\{a, b\} \mid a p b\} \in P(P_2(V))$

**relazione di adiacenza**: due estremi sono adiacenti se sono estremi dello stesso lato.

Due lati si dicono incidenti se hanno un vertice in comune.



Un vertice e un lato si dicono incidenti se il vertice è un estremo del lato.

Grado di un vertice:

$$\forall v \in V \quad \deg(v) = |\{l \in L \mid v \text{ è estremo di } l\}|$$

In altre parole, dipende dal numero di lati collegati a quell'estremo.

Poniamo  $G = (V, L)$ ,  $|V| = n \Rightarrow |L| \leq \binom{n}{2}$

$G$  si dice grafo completo quando  $L = P_2(V)$

Esempi:

$K_1$

$K_2$

$K_3$

$K_4$

$K_5$

etc.



Grafo planare: grafo in cui i lati sono rappresentabili senza intersezioni tra loro.



## Isomorfismo tra grafi (o multigrafi)

$$G = (V, L, f) \quad G' = (V', L', f')$$

isomorfismo di:  $G = G'$ :

$$\alpha: V \rightarrow V' \text{ biettiva} \quad \beta: L \rightarrow L'$$

tale che  $\forall l \in L \quad \forall a, b \in V$

$a$  e  $b$  sono gli estremi di  $l \Leftrightarrow \alpha(a)$  e  $\alpha(b)$  sono gli estremi di  $\beta(l)$ .

Per i grafi semplici basta:

$$G = (V, L) \quad G' = (V', L')$$

$\alpha: V \rightarrow V'$  biettiva tale che:

$$\forall a, b \in V \quad (\{a, b\} \in L \Leftrightarrow \{\alpha(a), \alpha(b)\} \in L')$$

ogni lato è definito univocamente da una coppia di vertici

**Teorema.** Sia  $G(V, L, f)$  un multigrafo finito.

$$\text{Allora } \sum_{v \in V} \deg(v) = 2|L| \quad \begin{array}{l} \text{si dimostra per induzione (o lati = grado o si aggiunge} \\ \text{lato raddoppiano gli estremi)} \end{array}$$

Contiamo in modo diverso due volte,  
sestremmo lo stesso risultato otteniamo  
un'ugualanza.

	$l_1$	$l_2$	$\dots$	$l_s$	$l_{s+1}$
$v_1$				x	
$v_2$					x
$\vdots$	x				
$v_k$					

Consideriamo  $S = \{(v, l) \in V \times L \mid v \text{ estremo di } l\}$   
Se contiamo per riga, ogni vertice avrà tante "x"  
quanti è il grado di  $v$ .

$$\text{Dunque } |S| = \sum_{v \in V} d(v)$$

Per colonne, invece, avendo ogni lato due vertici,  
avranno  $|S| = 2 \cdot |L|$ .

## Vertici pari e dispari

numero di grado pari  
arbitrario

di grado dispari  
arbitrario

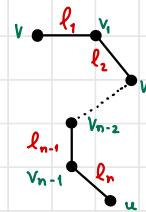
$\rightarrow$  non di numero arbitrario

Se  $\sum_{v \in V} d(v) = 2|E|$  in mod 2 quelli pari non incidono, quelli dispari devono essere di numero pari.

## Cammini

$\forall v, u \in V$  un cammino  $\gamma$  da  $v$  a  $u$  di lunghezza  $n \in \mathbb{N}$  è una  
 $n$ -pla  $(l_1, l_2, \dots, l_n) \in L^n$  di lati a due a due distinti tali che:

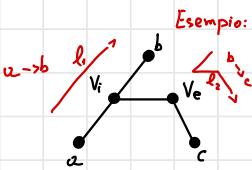
- $l_1$  ha estremi  $v$  ed un vertice  $v_1$
- $l_2$  ha estremi  $v_1$  e  $v_2$
- $\vdots$
- $l_{n-1}$  ha estremi  $v_{n-2}$  e  $v_{n-1}$
- $l_n$  ha estremi  $v_{n-1}$  e  $u$



$v$  e  $u$  si dicono connessi  $\Leftrightarrow \exists$  cammino tra  $v$  e  $u$

La relazione di connessione è riflessiva, simmetrica e transitiva.

$\rightarrow$  non si nota sempre dal



Esempio:  
Ma bisogna considerare il  
vertice  $v_i$  tale che:  
è il minimo tra gli indici dei  
vertici in comune tra i cammini

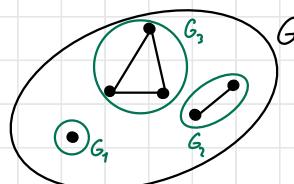
$a \rightarrow b \rightarrow c$  non è un cammino

ma  $a \rightarrow c$  è possibile considerando  $v_i$  come vertice in comune

La relazione di connessione è dunque d'equivalenza: possiamo definire classi di equivalenza e insieme quoziente.

Le classi sono dette componenti connesse e sono formate da vertici connessi e lati che formano i cammini che li connettono.

Se  $G$  è un grafo come rappresentato,  
la rel. di connessione,  $G_G = \{G_1, G_2, G_3\}$



## Sotto-grafi

Sia  $G = (V, L)$  un grafo.

Allora  $G' = (V', L')$  è un sottografo di  $G$  se e solo è un grafo,  $V' \subseteq V$ ,  $L' \subseteq L$ .

Per i multigrafi, è necessario che  $f'$  sia una riduzione di  $f$ .

## Distanza e Circuiti

Tra i cammini, il cammino più breve è detto distanza.

Un ciclo è un cammino da un vertice a se stesso.

Se un cammino (ciclo) passa per ogni lato di un multigrafo una ed una sola volta, allora si dice cammino (ciclo) euleriano.

N.B. "breve" := quello costituito dal minore numero di elementi dell'impla (quindi con  $l'n$  minore).

Condizioni necessarie e sufficienti:

- In un multigrafo finito e连通 esiste un circuito euleriano  $\Leftrightarrow$  ogni vertice ha grado pari.
- In un multigrafo finito e连通 esiste un cammino euleriano  $\Leftrightarrow$  ci sono esattamente 0 o 2 vertici di grado dispari. (il cammino collega i 2 vertici dispari)

## Foreste e Alberi

Un grafo è una foresta  $\Leftrightarrow$  è aciclico  $\Leftrightarrow$  è privo di circuiti non banali (di lunghezza 0).

Una foresta连通 si dice albero.

Teorema. Un grafo finito  $G$  è una foresta  $\Leftrightarrow$  per ogni coppia  $(a, b)$  con  $a \neq b$  di  $G$  esiste al più un cammino in  $G$  da  $a$  a  $b$ .

" " " è un albero  $\Leftrightarrow$  " " esiste esattamente un cammino in  $G$  da  $a$  a  $b$ .

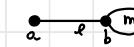
N.B. Una foresta è necessariamente un grafo semplice.

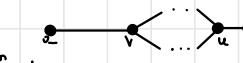
DIM Siano  $a, b \in V$ , dove  $V$  è l'insieme dei vertici della foresta  $G$

Supponiamo esistano due cammini diversi da  $a$  a  $b$ :

$l = (l_1, l_2, \dots, l_n)$  e  $m = (m_1, m_2, \dots, m_n)$

$\exists v \in V$ : da  $v$  in poi i due cammini  $l, m$  non coincidono.

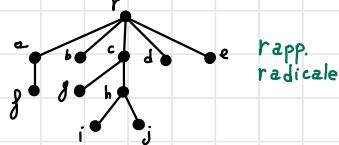
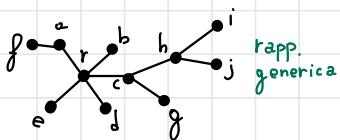
1.  $v = b$    $\Rightarrow$  non è un grafo aciclico, contraddizione.

2.  $v \neq b \Rightarrow$    $\exists u \in V$ : da  $u$  in poi i due cammini  $l, m$  coincidono

Prendendo in considerazione i cammini da  $v$  a  $u$  e da  $u$  a  $v$ , questi formano un circuito, contraddizione.

## Rappresentazione Radicale di un albero

Si può rappresentare un albero scegliendo un suo vertice  $r$  e disegnando al di sotto di esso quelli a distanza uno, poi due, etc. collegandoli per avere un grafo isomorfo.



OSS Una foresta finita è necessariamente un grafo plenare.

Se un albero ha almeno 2 vertici in rappresentazione radicale quelli a distanza massima hanno grado 1 (detti foglie). In generale il numero delle foglie è  $\geq d(r)$ .

Lemma. Sia  $T = (V, L)$  un albero,  $v \in V : d(v) = 1 \Rightarrow v$  è una foglia, sia  $l \in L : v$  è un estremo di  $l$ .

Allora  $T' = (V \setminus \{v\}, L \setminus \{l\})$  è ancora un albero.

DIM  $\forall a, b \in T' \quad a \neq v \wedge b \neq v \Rightarrow$  il cammino da  $a$  a  $b$  non è  $l$ , perché ogni vertice in un cammino ha grado  $\geq 2$  se non è un estremo.

Se  $T$  è finito,  $|V| = |L| + 1$ .

DIM Poniamo  $|L| = 0 \Rightarrow$  quindi  $|V| = 1$ . ✓

Assumiamo  $|L| = n \Rightarrow$  verifichiamo  $|L_{f_i}| = n+1$

Per il lemma precedente, se a  $t'$  togliamo una foglia e il lato ad essa incidente, abbiamo ancora un albero di  $n+1-1 = n$  lati e  $|V_{f_i}| - 1$  vertici.

Per ipotesi i vertici di quest'albero sono  $n+1$ , dunque  $|V_{f_i}| - 1 = n+1 \Rightarrow |V_{f_i}| = n+1+1 = |L_{f_i}| + 1$ .

## Sottoalberi Massimali (o alberi di supporto)

Sia  $G = (V, L, f)$  un multigrafo finito.

Un sottoalbero massimale di  $G$  è un sottografo  $G'$  con insieme di vertici  $V'$  che sia un albero.

È unico  $\Leftrightarrow G$  è连通的.

Si dimostra facilmente che se eliminano un lato da un circuito  $i : v \in V$  rimangono tutti connnessi.

OSS Sia  $G = (V, L, f)$  un multigrafo connesso. Sia  $l \in L$ :  $l$  è parte di un circuito in  $G$ .

Allora il sottografo di  $G$ ,  $G' = (V, L', f')$  con  $L' = L \setminus \{l\}$  è ancora connesso.

DIM Siano  $a, b \in G$ :

$a, b$  connessi da un circuito  $\Rightarrow$  esistono almeno due cammini diversi ( $n \neq m$ ) da  $a \rightarrow b$ .

Se  $l_0$  è in  $n \Rightarrow a \text{ e } b$  sono connessi da  $m$  in  $G'$ , o viceversa.

## Teorema conclusivo.

Sia  $G = (V, L, f)$  multigrado finito con esattamente  $k$  componenti connesse. Allora:

1.  $G$  connesso  $\Rightarrow |L| \geq |V| - k$

2.  $|L| \geq |V| - k$

3.  $|L| = |V| - k \iff G$  è una foresta

4. Sono equivalenti: a.  $G$  è un albero; b.  $G$  è connesso e  $|V| = |L| + 1$ ; c.  $G$  è una foresta e  $|V| = |L| + 1$

DIM

1.  $G$  ha un sottoalbero massimale  $= (V, L')$  con  $L' \subseteq L$ .

Essendo un albero, allora  $|V| = |L'| + 1$ . Quindi  $|L'| = |V| - 1$ , ed essendo  $L' \subseteq L$ ,  $|L'| \leq |L|$ .  
 $\Rightarrow |L| \geq |V| - k$ .

Vale l'uguaglianza quando  $L' = L$ , dunque  $G$  è il sottoalbero massimale di sé.

4. (b  $\Rightarrow$  a) per quanto appena detto. (a  $\Rightarrow$  b) già dimostrato.

2. Siano  $(V_1, L_1, f_1), (V_2, L_2, f_2), \dots, (V_k, L_k, f_k)$  le componenti connesse di  $G$ .

$$\forall i \in \{1, 2, \dots, k\}, |L_i| \geq |V_i| - 1 \text{ per } 1.$$

$$|L| = \sum_{i=1}^k |L_i|, |V| = \sum_{i=1}^k |V_i|, \sum_{i=1}^k -1 = -k$$

Dunque  $|L| \geq |V| - k$ .

3. Se  $\forall i \in \{1, 2, \dots, k\}, |L_i| = |V_i| - 1$ , allora otteniamo che  $|L| = |V| - k$  per 2.

Inoltre  $G_i = (V_i, L_i)$  è un albero, dunque  $G$  è una foresta.

4. (a  $\Rightarrow$  c) Se  $G$  è un albero, allora  $G$  è anche una foresta. Inoltre per 3  $|V| = |L| + k$ , dove  $k$  sono le componenti connesse in  $G$ . Essendo  $G$  un albero,  $k = 1$ .

(c  $\Rightarrow$  a) per 3  $G$  è una foresta con  $k = 1$  componenti connesse, dunque un albero.