5756-Gherardi.qxp_SN3 11/01/23 13:28 Page I

Studi di epistemologia

Collana diretta da Maria Carla Galavotti

6



Le opere pubblicate nella collana sono sottoposte all’approvazione
di un rappresentante del comitato scientifico e di due componenti esterni.

5756-Gherardi.qxp_SN3  11/01/23  13:28  Page III

Guido Gherardi

# Introduzione alla Teoria dei Modelli

archetipolibri

Questo volume è stato finanziato con il contributo del Dipartimento di Filosofia e Comunicazione dell'Alma Mater Studiorum – Università di Bologna.


Copertina e progetto grafico: Avenida (Modena)

ISBN 978-88-6633-179-7

ArchetipoLibri
40126 Bologna - Via Marsala 31
www.archetipolibri.it / www.clueb.it
ArchetipoLibri è un marchio Clueb

Finito di stampare nel mese di gennaio 2023
da Editografica - Rastignano (Bo)

# Indice

# Capitolo 1

# Introduzione

Scopo di questo manuale è quello di tracciare alcuni argomenti di base per un'agile introduzione allo studio della Teoria dei Modelli, l'ambito della Logica che ad oggi forse si è rivelato essere il più proficuo nelle applicazioni alla Matematica tradizionale. Il suo contenuto è pensato primariamente quale strumento di supporto per gli studenti del mio corso di Teoria di Modelli annualmente tenuto presso l'Università di Bologna, e frequentato principalmente da studenti di Matematica e di Filosofia, ma è anche concepito per essere adatto, auspicabilmente, per lo studio individuale da parte di chiunque possa essere interessato alla disciplina, essendo provvisto di nozioni matematiche di livello scolastico. Sulla base di quelle, questo testo cercherà di fornire i rudimenti di quegli argomenti di Matematica a cui verranno poi applicati i metodi della Teoria dei Modelli. In particolare, verranno fornite nozioni elementari di Algebra e Topologia. Tuttavia, questa parte introduttiva non si limiterà a riassumere concetti che probabilmente saranno già noti agli studenti di Matematica, ma presenterà anche le principali strutture algebriche astratte (monoidi, gruppi, anelli e campi) da un punto di vista assiomatico formale espresso nel linguaggio logico al primordine. Non sono stati esplicitamente definiti in questa parte alcuni concetti di Analisi Reale, come la nozione di convergenza di una successione o quella di successione di Cauchy, che verranno ricordati direttamente laddove saranno utilizzati (nel caso specifico, nel capitolo sull'Analisi Non Standard). Verranno fornite brevemente anche nozioni basilari di Logica, seppur in modo più succinto, rimandando i lettori a manuali d'introduzione alla Logica al primordine per trattazioni più approfondite.

Gli argomenti di Teoria dei Modelli selezionati in questo volume sono il Teorema di Compattezza, le teorie assiomatizzabili universalmente, la categoricità, l'eliminazione dei quantificatori, la decidibilità delle teorie, gli ultrafiltri, l'Analisi Non Standard.

Questo libro è dedicato alla memoria di Franco Montagna, a cui devo l'apprendimento delle basi della Teoria dei Modelli. Un ringraziamento speciale va ad Alessandro Cobbe, per le sue consulenze algebriche di fondamentale importanza. Ringrazio anche alcuni miei ex-studenti che hanno trovato refusi nelle originarie dispense da cui il volume è stato tratto o che mi hanno dato utili informazioni matematiche: Miriam Abbate, Melissa Antonelli, Giuseppe Bianco, Davide Davoli, Enrico Maresca, Riccardo Zanichelli. In particolare ringrazio accoratamente il mio ex-studente Nicola Carissimi per avermi introdotto alla dimostrazione topologica del Teorema di Compattezza Logica. Non posso inoltre astenermi dal ringraziare Giovanna Corsi per avere occupato un posto insostituibile nella mia formazione logica e per avermi dato suggerimenti essenziali per il capitolo sull'introduzione alla Logica. Un ringraziamento va ancora ad Eugenio Orlandelli per il suo supporto nella redazione finale del *file latex* del libro. Sono infine riconoscente a Maria Carla Galavotti per avere accolto con favore la pubblicazione del mio libro nella collana "Studi di Epistemologia" da lei curata.

# Capitolo 2

# Basi logiche

## 2.1 Semantica per la Logica del primordine con identità

Utilizzeremo la logica al primordine con l'identità, avente il simbolo di uguaglianza come costante logica. Diamo per scontata la familiarità con i concetti base della sintassi del linguaggio al primordine con identità. Per maggiori dettagli rimandiamo a Mendelson [1972].

Un linguaggio $\mathscr{L}$ al primordine è costituito da simboli di funzione (compresi i simboli di costante) e da simboli di predicato non logici.

Essendo = un simbolo logico, non lo includeremo nell'insieme dei predicati in $\mathscr{L}$.

Una $\mathscr{L}$-struttura algebrica $\mathscr{A}$ è data da un insieme $A \neq \emptyset$ (detto *dominio* della struttura) e da funzioni di interpretazione su $A$ per ogni simbolo funzionale e predicativo del linguaggio $\mathscr{L}$. Più precisamente, ad ogni simbolo di costante $c$ viene associato un individuo $c^{\mathscr{A}} \in A$, ad ogni simbolo di funzione $n$-ario $f^n$ viene associata una funzione $(f^n)^{\mathscr{A}} : A^n \to A$, e ad ogni simbolo di predicato $n$-ario $P^n$ un predicato $(P^n)^{\mathscr{A}} \subseteq A^n$, per $n \geq 1$.

N.B.: per semplicità, in molti casi concreti non distingueremo notazionalmente tra i simboli del linguaggio e le loro rispettive interpretazioni, in accordo con l'usuale pratica matematica. Ciò vale già per il simbolo di uguaglianza e la relazione di identità che esso denota. Inoltre sovente tralasceremo di scrivere l'arietà nei simboli di funzione e di predicato.

Consideriamo una $\mathscr{L}$-struttura $\mathscr{A}$ sull'insieme $A$. Per $t(x_1,..,x_n)$ qualsiasi $\mathscr{L}$-termine le cui variabili (se presenti) siano in $\{x_1,...,x_n\}$ ($x_i \neq x_j$ per $i \neq j$), definiamo $t^{\mathscr{A}} : A^n \to A$ per induzione sulla costruzione dei termini. Più precisamente, siano $a_1,...,a_n \in A$. Definiamo allora per induzione sulla costruzione dei termini l'individuo $t^{\mathscr{A}}(a_1,...,a_n)$ nel seguente modo[1]:

- $t(x_1,..,x_n) \equiv c$, per $c$ costante, allora $t^{\mathscr{A}}(a_1,...,a_n) := c^{\mathscr{A}} \in A$,
- $t(x_1,..,x_n) \equiv x_i$, per $i \in \{1,...,n\}$, allora $t^{\mathscr{A}}(a_1,...,a_n) := a_i \in A$,
- $t(x_1,..,x_n) \equiv f(t_1(x_1,...,x_n),...,t_m(x_1,...,x_n))$,
  allora $t^{\mathscr{A}}(a_1,...,a_n) := f^{\mathscr{A}}(t_1^{\mathscr{A}}(a_1,..,a_n),...,t_m^{\mathscr{A}}(a_1,...,a_n)) \in A$.

Per $\varphi(x_1,...,x_n)$ $\mathscr{L}$-formula con variabili libere[2] (possibilmente nessuna) in $\{x_1,...,x_n\}$ ($x_i \neq x_j$ per $i \neq j$), definiamo

$$\mathscr{A} \models \varphi(x_1,...,x_n)[a_1,...,a_n]$$

con $a_1,..,a_n \in A$ (dicendo che $\varphi(x_1,...,x_n)$ *è soddisfatta in* $\mathscr{A}$ *per* $a_1$, ..., $a_n$) per induzione sulla costruzione delle formule:

- $\varphi(x_1,...,x_n) \equiv P(t_1(x_1,...,x_n),..,t_m(x_1,...,x_n))$: allora
  $$\mathscr{A} \models P(t_1(x_1,...,x_n),..,t_m(x_1,...,x_n))[a_1,..,a_n]$$
  qualora $(t_1^{\mathscr{A}}(a_1,..,a_n),...,t_m^{\mathscr{A}}(a_1,..,a_n)) \in P^{\mathscr{A}} \subseteq A^m$
- $\varphi(x_1,...,x_n) \equiv t_1(x_1,...,x_n) = t_2(x_1,...,x_n)$: allora
  $$\mathscr{A} \models t_1(x_1,...,x_n) = t_2(x_1,...,x_n)[a_1,...,a_n]$$
  qualora $t_1^{\mathscr{A}}(a_1,..,a_n) = t_2^{\mathscr{A}}(a_1,..,a_n)$, ovvero $t_1^{\mathscr{A}}(a_1,..,a_n)$ e $t_2^{\mathscr{A}}(a_1,..,a_n)$ siano lo stesso individuo

[1] Utilizzeremo il simbolo $\equiv$ per indicare l'identità sintattica tra oggetti linguisitici.

[2] Sono libere quelle variabili con alcune occorrenze non vincolate da quantificatori. Ad esempio in $(\exists x)(\forall y)(P(x,z) \to Q(y)) \wedge R(x)$ le variabili libere sono $x$ e $z$ (si noti che $z$ è libera semplicemente in quanto non vi è nessun quantificatore $(\forall z)$, $(\exists z)$ nella formula, mentre $x$ è libera in quanto sebbene nella formula vi sia il quantificatore $(\exists x)$ esso non agisce sull'occorrenza più a destra di $x$).

- $\varphi(x_1,...,x_n) \equiv \perp$: allora
$$\mathscr{A} \not\models \perp(x_1,...,x_n)[a_1,..,a_n]$$
- $\varphi(x_1,...,x_n) \equiv \neg\psi(x_1,..,x_n)$: allora
$$\mathscr{A} \models \neg\psi(x_1,...,x_n)[a_1,..,a_n]$$
qualora $\mathscr{A} \not\models \psi(x_1,...,x_n)[a_1,..,a_n]$
- $\varphi(x_1,...,x_n) \equiv \psi(x_1,...,x_n) \land \xi(x_1,...,x_n)$: allora
$$\mathscr{A} \models \psi(x_1,...,x_n) \land \xi(x_1,...,x_n)[a_1,..,a_n]$$
qualora $\mathscr{A} \models \psi(x_1,...,x_n)[a_1,...,a_n]$ e $\mathscr{A} \models \xi(x_1,...,x_n)[a_1,...,a_n]$
- $\varphi(x_1,...,x_n) \equiv \psi(x_1,...,x_n) \lor \xi(x_1,...,x_n)$: allora
$$\mathscr{A} \models \psi(x_1,...,x_n) \lor \xi(x_1,...,x_n)[a_1,...,a_n]$$
qualora $\mathscr{A} \models \psi(x_1,...,x_n)[a_1,...,a_n]$ o $\mathscr{A} \models \xi(x_1,...,x_n)[a_1,...,a_n]$
- $\varphi(x_1,...,x_n) \equiv \psi(x_1,...,x_n) \to \xi(x_1,...,x_n)$: allora
$$\mathscr{A} \models \psi(x_1,...,x_n) \to \xi(x_1,...,x_n)[a_1,...,a_n]$$
qualora $\mathscr{A} \not\models \psi(x_1,...,x_n)[a_1,...,a_n]$ o $\mathscr{A} \models \xi(x_1,...,x_n)[a_1,...,a_n]$
- $\varphi(x_1,...,x_n) \equiv \psi(x_1,...,x_n) \leftrightarrow \xi(x_1,...,x_n)$: allora
$$\mathscr{A} \models \psi(x_1,...,x_n) \leftrightarrow \xi(x_1,...,x_n)[a_1,...,a_n]$$
qualora $\mathscr{A} \models \psi(x_1,...,x_n)[a_1,...,a_n]$ sse $\mathscr{A} \models \xi(x_1,...,x_n)[a_1,...,a_n]$
- $\varphi(x_1,...,x_n) \equiv (\forall x)\psi(x,x_1,...,x_n)$: allora
$$\mathscr{A} \models (\forall x)\psi(x,x_1,...,x_n)[a_1,...,a_n]$$
qualora per ogni $a \in A$ si dia $\mathscr{A} \models \psi(x,x_1,...,x_n)[a,a_1,...,a_n]$[3]

[3]Affinché la definizione sia consistente, bisognerebbe formularla rispetto all'indice delle (meta)variabili nel seguente modo: $\mathscr{A} \models (\forall x_i)\psi(x_1,...,x_i,...,x_n)[a_1,...,a_n]$ qualora per ogni $a \in A$ si dia $\mathscr{A} \models \psi(x_1,...,x_i,...,x_n)[a_1,...,a_{i-1},a,a_{i+1},...,a_n]$ (vedasi Chang & Keisler [1973]). Tuttavia per semplicità utilizzeremo la notazione semplificata data sopra, simile a quella di Marcja & Toffalori [2003] o Marker [2002] (che però presuppone restrizioni nell'uso delle variabili nella costruzione delle formule). Un discorso analogo vale per il caso del quantificatore esistenziale $\exists$.

- $\varphi(x_1,\ldots,x_n) \equiv (\exists x)\psi(x,x_1,\ldots,x_n)$: allora

$$\mathscr{A} \models (\exists x)\psi(x,x_1,\ldots,x_n)[a_1,\ldots,a_n]$$

  qualora esista un $a \in A$ tale che $\mathscr{A} \models \psi(x,x_1,\ldots,x_n)[a,a_1,\ldots,a_n]$

Qualora esistano $a_1,\ldots,a_n \in A$ tali che $\mathscr{A} \models \varphi(x_1,\ldots,x_n)[a_1,\ldots,a_n]$ allora diciamo che $\varphi(x_1,\ldots,x_n)$ è *soddisfacibile* in $\mathscr{A}$.

Spesso scriveremo più semplicemente $\mathscr{A} \models \varphi[a_1,\ldots,a_n]$ lasciando per implicite le variabili per le quali viene eseguita l'interpretazione in $A$ (questo varrà in particolare per ovvie ragioni per le formule chiuse, ovvero quelle formule in cui non vi sono variabili libere).

**Definizione 2.1** (Modelli)**.** *Qualora valga $\mathscr{A} \models \varphi[a_1,\ldots,a_n]$ per tutti gli $a_1,..,a_n \in A$, diciamo che $\mathscr{A}$ è un* modello *per $\varphi(x_1,\ldots,x_n)$ e che $\varphi(x_1,\ldots,x_n)$ è* vera in *$\mathscr{A}$; in tal caso scriviamo $\mathscr{A} \models \varphi(x_1,\ldots,x_n)$, o anche semplicemente $\mathscr{A} \models \varphi$ qualora $\varphi$ sia chiusa o qualora non si vogliano specificare le sue possibili variabili libere.*

*Per $\Gamma$ insieme di $\mathscr{L}$-formule ed $\mathscr{A}$ $\mathscr{L}$-struttura scriviamo $\mathscr{A} \models \Gamma$ quando $\mathscr{A} \models \varphi$ per ogni $\varphi \in \Gamma$. In tale caso diciamo che $\mathscr{A}$ è un* modello *per* $\Gamma$.

*Per una $\mathscr{L}$-formula $\varphi(x_1,..,x_n)$ scriviamo $\models \varphi(x_1,..,x_n)$ (o anche semplicemente $\models \varphi$ qualora $\varphi$ sia chiusa o non si vogliano specificare le sue possibili variabili libere) se $\mathscr{A} \models \varphi(x_1,\ldots,x_n)$ per ogni $\mathscr{L}$-struttura $\mathscr{A}$. In questo caso diciamo che $\varphi(x_1,\ldots,x_n)$ è una* formula valida *o* verità logica *della logica al primordine con identità nel linguaggio $\mathscr{L}$.*

**Osservazione 2.2.** *Una formula chiusa $\varphi$ è soddisfatta in $\mathscr{A}$ per alcuni $a_1,\ldots,a_n \in A$ sse è soddisfatta per tutti gli $a_1,..,a_n \in A$. In altre parole una formula chiusa $\varphi$ è soddisfacibile in $\mathscr{A}$ sse $\varphi$ è vera in $\mathscr{A}$ sse $\mathscr{A}$ è modello di $\varphi$.*

*Per questa ragione diremo equivalentemente che un insieme di formule chiuse $\Gamma$ è soddisfacibile oppure che ha un modello.*

**Definizione 2.3** (Conseguenza Logica ed Equivalenza)**.** *Dati un insieme di formule $\Gamma$ ed una formula $\varphi$ nel linguaggio $\mathscr{L}$, diciamo che $\varphi$ è* conseguenza logica *di $\Gamma$ ($\Gamma \models \varphi$) se per ogni $\mathscr{L}$-struttura $\mathscr{A}$:*

$$\mathscr{A} \models \Gamma \Rightarrow \mathscr{A} \models \varphi.$$

*Dati insiemi di formule* $\Gamma, \Delta$ *nel linguaggio* $\mathscr{L}$, *diciamo che* $\Delta$ *è conseguenza logica di* $\Gamma$ ($\Gamma \models \Delta$) *se* $\Gamma \models \varphi$ *per ogni* $\varphi \in \Delta$.

*Due* $\mathscr{L}$*-formule* $\varphi(x_1,\dots,x_n)$ *e* $\psi(x_1,\dots,x_n)$ *sono* logicamente equivalenti*, notazionalmente*

$$\varphi(x_1,\dots,x_n) \approx \psi(x_1,\dots,x_n),$$

*se* $\varphi(x_1,\dots,x_n) \models \psi(x_1,\dots,x_n)$ *e* $\psi(x_1,\dots,x_n) \models \varphi(x_1,\dots,x_n)$.

**Osservazione 2.4.** *Ciascuna formula aperta* $\varphi(x_1,\dots,x_n)$ *è logicamente equivalente alla sua chiusura universale, cioè*

$$\varphi(x_1,\dots,x_n) \approx (\forall x_1)\dots(\forall x_n)\varphi(x_1,\dots,x_n).$$

*Si noti che* $\models \varphi \leftrightarrow \psi$, *per* $\varphi$ *o* $\psi$ *eventualmente contenenti variabili libere, implica* $\varphi \approx \psi$, *mentre l'inverso in generale non vale. Si pensi ad esempio al caso in cui* $\varphi$ *abbia delle variabili libere e* $\psi$ *sia la sua chiusura universale. Tuttavia qualora sia* $\varphi$ *che* $\psi$ *siano formule chiuse anche l'inverso vale.*

## *2.2 Calcoli alla Hilbert per la Logica al primordine con identità*

Dato un linguaggio $\mathscr{L}$ al primordine, i seguenti schemi di formule sono detti *assiomi* della logica al primordine con identità (nel linguaggio $\mathscr{L}$):

(P1) $\varphi \to (\psi \to \varphi)$ (*a fortiori*)

(P2) $(\varphi \to (\psi \to \chi)) \to ((\varphi \to \psi) \to (\varphi \to \chi))$ (Frege)

(P3) $(\varphi \to \psi) \to ((\varphi \to \neg\psi) \to \neg\varphi)$ (assurdo 1)

(P4) $(\neg\varphi \to \psi) \to ((\neg\varphi \to \neg\psi) \to \varphi)$ (assurdo 2)

(P5) $\perp \to \varphi$ (*ex falso quodlibet*)

(P6) $\varphi \to (\neg\varphi \to \perp)$ (non contraddizione)

(P7) $(\varphi \to \psi) \to ((\varphi \to \chi) \to (\varphi \to \psi \wedge \chi))$ (introduzione della congiunzione)

(P8) $\varphi \wedge \psi \to \varphi$ (eliminazione della congiunzione 1)

(P9) $\varphi \wedge \psi \to \psi$ (eliminazione della congiunzione 2)

(P10) $\varphi \to \varphi \vee \psi$ (introduzione della disgiunzione 1)

(P11) $\psi \to \varphi \vee \psi$ (introduzione della disgiunzione 2)

(P12) $(\varphi \to \chi) \to ((\psi \to \chi) \to (\varphi \vee \psi \to \chi))$ (eliminazione della disgiunzione)

(P13) $(\varphi \leftrightarrow \psi) \to (\varphi \to \psi) \wedge (\psi \to \varphi)$ (doppia implicazione 1)

(P14) $(\varphi \to \psi) \wedge (\psi \to \varphi) \to (\varphi \leftrightarrow \psi)$ (doppia implicazione 2)

(Q1) $(\forall x)\varphi \to \varphi(t/x)$ per $t$ termine non contenente variabili vincolate in $\varphi$ (*dictum de omni*)

(Q2) $\varphi(t/x) \to (\exists x)\varphi$ per $t$ termine non contenente variabili vincolate in $\varphi$ (non vacuità)

(I1) $(\forall x)\, x = x$ (riflessività dell'identità)

(I2) $(\forall x)(\forall y)(x = y \to y = x)$ (simmetria dell'identità)

(I3) $(\forall x)(\forall y)(\forall z)(x = y \to (y = z \to x = z))$ (transitività dell'identità)

(I4) $(\forall x_1)\dots(\forall x_n)(\forall y_1)\dots(\forall y_n)(x_1 = y_1 \wedge \dots \wedge x_n = y_n \to f^n(x_1,\dots,x_n) = f^n(y_1,\dots,y_n))$ per ogni simbolo di funzione $f^n$ con $n \geq 1$ (definizione di funzione)

(I5) $(\forall x_1)\dots(\forall x_n)(\forall y_1)\dots(\forall y_n)(x_1 = y_1 \wedge \dots \wedge x_n = y_n \to (P^n(x_1,\dots,x_n) \to P^n(y_1,\dots,y_n))$ per ogni simbolo di predicato $P^n$ con $n \geq 1$ (principio di estensionalità)

**Definizione 2.5** (Deduzione)**.** *Dato un insieme di $\mathscr{L}$-formule $\Gamma$, diciamo che la successione $(\varphi_1,\dots,\varphi_n)$ di $\mathscr{L}$-formule è una* deduzione *(o* derivazione*) di $\varphi_n$ da $\Gamma$ se ciascuna $\varphi_i$, per $1 \leq i \leq n$, è una delle seguenti: i) elemento di $\Gamma$, ii) assioma del calcolo al primordine con identità, iii) ottenuta da due formule $\varphi_j, \varphi_k$, per $j, k < i$, tramite Modus Ponens (MP), iv) ottenuta da una formula $\varphi_j$ per $j < i$ tramite Generalizzazione (Gen) o Particolarizzazione (Part), dove*

$$\frac{\varphi \to \psi \quad \varphi}{\psi}\ MP$$

$$\frac{\varphi \to \psi}{\varphi \to (\forall x)\,\psi}\ Gen \qquad\qquad x \text{ non libera in } \varphi$$

$$\frac{\varphi \to \psi}{(\exists x)\,\varphi \to \psi}\ Part \qquad\qquad x \text{ non libera in } \psi$$

*In tal caso diciamo che* $\varphi_n$ *è* deducibile *(o* derivabile*) da* $\Gamma$ *e scriviamo* $\Gamma \vdash \varphi_n$.

*Per* $\Gamma, \Delta$ *insiemi di* $\mathscr{L}$*-formule scriviamo* $\Gamma \vdash \Delta$ *se* $\Gamma \vdash \varphi$ *per ogni* $\varphi \in \Delta$.

**Definizione 2.6** (Consistenza)**.** *Un insieme di* $\mathscr{L}$*-formule* $\Gamma$ *è detto* consistente *se* $\Gamma \nvdash \bot$ *(equivalentemente, se* $\Gamma \nvdash \varphi$ *per almeno una* $\mathscr{L}$*-formula* $\varphi$*).*

Com'è ben noto, il calcolo al primordine è sia valido che completo:

**Teorema 2.7** (Validità)**.** $\Gamma \vdash \varphi \Rightarrow \Gamma \models \varphi$.

*In generale,* $\Gamma \vdash \Delta \Rightarrow \Gamma \models \Delta$.

**Teorema 2.8** (Completezza)**.** $\Gamma \models \varphi \Rightarrow \Gamma \vdash \varphi$.

*In generale,* $\Gamma \models \Delta \Rightarrow \Gamma \vdash \Delta$.

**Definizione 2.9** (Teoria)**.** *Un insieme di* $\mathscr{L}$*-formule* $\Gamma$ *è* deduttivamente chiuso *se per ogni* $\mathscr{L}$*-formula* $\varphi$*,* $\Gamma \vdash \varphi \Rightarrow \varphi \in \Gamma$.

*Una* teoria $T$ *è un insieme di formule chiuse deduttivamente chiuso.*

$T$ *è una teoria per* $\mathscr{A}$ *se* $\mathscr{A} \models T$*, ed è una teoria per una classe* $\mathscr{C}$ *di strutture se* $\mathscr{A} \models T$ *per ogni* $\mathscr{A} \in \mathscr{C}$.

*Dato un sottoinsieme* $\Gamma \subseteq T$*, diciamo che* $\Gamma$ *è un* sistema di assiomi *per* $T$ *se* $\Gamma \vdash T$.

Ovviamente l'intera $T$ è un sistema di assiomi per sé stessa, ma ovviamente è un sistema d'assiomi poco interessante. Per fortuna vi sono teorie importanti che ammettono sistemi di assiomi interessanti. Ne vedremo diversi esempi nel seguente capitolo.

Capitolo 3

# Elementi di Algebra e Topologia

## 3.1 Importanti strutture algebriche

Ricordiamo ora alcuni tipi fondamentali di strutture algebriche riportando sistemi di assiomi che li caratterizzano. Intendiamo con ciò dire che la teoria $T$ delle strutture algebriche di un certo tipo, rispetto ad uno specifico linguaggio $\mathscr{L}$, è assiomatizzata dagli assiomi corrispondenti che daremo qui di seguito[1].

- **Monoidi:** un *monoide* nel linguaggio $\mathscr{L} := \{\dot{+}\}$ è una coppia $\mathscr{M} := \langle M, \dot{+}\rangle$, per $M$ insieme non vuoto e $\dot{+} : M \times M \to M$ operazione binaria su $M$, tale che

  (m1) $(\forall x)(\forall y)(\forall z)\, x \dot{+} (y \dot{+} z) = (x \dot{+} y) \dot{+} z$ (associatività)

  (m2) $(\exists y)(\forall x)\, x \dot{+} y = x = y \dot{+} x$ (elemento neutro)

  Si può dimostrare che l'elemento neutro di un monoide è univocamente determinato. Siano infatti $y$ ed $y'$ elementi neutri di $\mathscr{M}$. Si ha allora che $y = y \dot{+} y' = y'$.

[1] Per semplicità e in analogia con l'usuale pratica matematica, per i tipi di strutture algebriche fondamentali che considereremo, non distingueremo notazionalmente tra i simboli del linguaggio e la loro interpretazione all'interno della struttura.

Questo fatto può suggerirci, in modo del tutto naturale, l'introduzione nel linguaggio di una costante $\dot{0}$ designata a denotare l'*unico* elemento neutro del monoide:

Un monoide nel linguaggio $\mathscr{L} := \{\dot{0}, \dot{+}\}$ è una tripla $\mathscr{M} := \langle M, \dot{0}, \dot{+}\rangle$, con $\dot{0} \in M$ e $\dot{+}$ funzione binaria, che soddisfa l'assioma (m1) e

(m2') $(\forall x)\, x \dot{+} \dot{0} = x = \dot{0} \dot{+} x$ (elemento neutro)

*Esempi:*

1. $\langle \mathbb{N}, 0, \max^2\rangle$, con $\max^2$ la funzione che restituisce il massimo di due elementi
2. $\langle \mathbb{N}, 0, +\rangle$
3. l'insieme delle stringhe finite su un alfabeto $\Sigma$ con stringa vuota (stringa di lunghezza 0) ed operazione di concatenazione

- **Gruppi:** un *gruppo* nel linguaggio $\mathscr{L} := \{\dot{0}, \dot{+}\}$ è una tripla $\mathscr{G} := \langle G, \dot{0}, \dot{+}\rangle$ tale che $\langle G, \dot{0}, \dot{+}\rangle$ è un monoide ed inoltre

  (g) $(\forall x)(\exists y)\, x \dot{+} y = \dot{0} = y \dot{+} x$ (elementi inversi)

  Se vale anche

  (gc) $(\forall x)(\forall y)\, x \dot{+} y = y \dot{+} x$ (commutatività)

  il gruppo è detto *commutativo* o *abeliano*.

  Si può dimostrare che per ogni $x$ l'elemento inverso di $x$ è univocamente determinato. Siano infatti $y, y'$ elementi inversi di $x$: allora $y = y \dot{+} \dot{0} = y \dot{+} (x \dot{+} y') = (y \dot{+} x) \dot{+} y' = \dot{0} \dot{+} y' = y'$.

  Questo fatto ci suggerisce analogamente la riformulazione del concetto di gruppo in un linguaggio che contenga un simbolo per un'operazione *unaria* $\dot{-}$ inversa rispetto a $\dot{+}$:

  Un gruppo nel linguaggio $\mathscr{L} := \{\dot{0}, \dot{-}, \dot{+}\}$ è una quadrupla $\mathscr{G} := \langle G, \dot{0}, \dot{-}, \dot{+}\rangle$ tale che $\langle G, \dot{0}, \dot{+}\rangle$ è un monoide e $\dot{-} : G \to G$ è un'operazione unaria tale che:

(g') $(\forall x)\, x \dotplus (\dot{-}\, x) = \dot{0} = \dot{-}\, x \dotplus x$ (elementi inversi)

Per ogni $x, y \in G$ scriveremo con abuso di linguaggio $x \dot{-}\, y$ al posto di $x \dotplus (\dot{-}\, y)$ in analogia con la notazione matematica per $+$ e $-$ ma non si deve dimenticare che intendiamo $\dot{-}$ come operazione *unaria* e non binaria!

Si osservi non solo che dato un qualsiasi oggetto in un gruppo il suo elemento inverso è univocamente determinato, ma anche che non esistono due oggetti distinti con il medesimo elemento inverso (la funzione di inverso è quindi iniettiva). Sia infatti $\dot{-}\, x = \dot{-}\, y$ per $x, y \in G$. Allora $x \dot{-}\, y = \dot{0}$ da cui $x = x \dotplus \dot{0} = x \dotplus (\dot{-}\, y \dotplus y) = (x \dot{-}\, y) \dotplus y = \dot{0} \dotplus y = y$. Da questo segue che se $x \neq y$ allora $x \dot{-}\, y \neq \dot{0}$.

In un gruppo si può anche dimostrare non solo che vi è un unico elemento neutro $\dot{0}$ che vada bene per tutti gli $x \in G$, ma anche che non vi è alcun $x$ che abbia più di un "elemento neutro relativamente ad esso", intendendo con questo più di un $y$ tale che $x \dotplus y = y \dotplus x = x$[2]. Sia infatti $y$ siffatto rispetto ad $x$. Allora $x \dotplus y = x = x \dotplus \dot{0}$, da cui $\dot{-}\, x \dotplus x \dotplus y = \dot{-}\, x \dotplus x \dotplus \dot{0}$, e quindi $y = \dot{0}$.

Il concetto di gruppo abeliano nel linguaggio $\mathscr{L} := \{\dot{0}, \dot{-}, \dotplus\}$ è definito di conseguenza.

*Esempi:*

1. $\langle \mathbb{Z}, 0, -, + \rangle$ è un gruppo abeliano
2. $\langle \mathbb{Q} \setminus \{0\}, 1, ^{-1}, \cdot \rangle$ con $(\frac{a}{b})^{-1} = \frac{b}{a}$ per $a, b \in \mathbb{Z} \setminus \{0\}$ è un gruppo abeliano
3. l'insieme delle permutazioni di $n$ elementi è un gruppo, non abeliano per $n > 2$

I primi due esempi considerati sono gruppi abeliani, il terzo no.

- **Anelli:** un *anello* nel linguaggio $\mathscr{L} := \{\dot{0}, \dot{-}, \dotplus, *\}$ è una quintupla $\mathscr{R} := \langle R, \dot{0}, \dot{-}, \dotplus, * \rangle$, con $\langle R, \dot{0}, \dot{-}, \dotplus \rangle$ gruppo abeliano e $* : R \times R \to R$ operazione binaria tale che:

[2] Al contrario, questo si dà nel caso già visto del monoide $\langle \mathbb{N}, 0, \max^2 \rangle$.

(r1) $(\forall x)(\forall y)(\forall z)\, x * (y * z) = (x * y) * z$ (associatività)

(r2) $(\forall x)(\forall y)(\forall z)\, x * (y \dotplus z) = x * y \dotplus x * z$ (distributività a sinistra)[3]

(r3) $(\forall x)(\forall y)(\forall z)\, (y \dotplus z) * x = y * x \dotplus z * x$ (distributività a destra)

Se vale anche

(rc) $(\forall x)(\forall y)\, x * y = y * x$ (commutatività)

l'anello è detto *commutativo*.

*Esempi:*

1. $\langle \mathbb{Z}, 0, -, +, \cdot \rangle$
2. $\langle \mathbb{Z}_k, 0, \hat{-}, + \mod k, \cdot \mod k \rangle$, dove $\mathbb{Z}_k := \{0, 1, \dots, k-1\}$ per $k \geq 2$, e $\hat{-}m := k - m$ per $1 \leq m \leq k-1$, $m + n \mod k := Resto(\frac{m+n}{k})$, $m \cdot n \mod k := Resto(\frac{m \cdot n}{k})$
3. L'insieme $\mathbb{Z}[x]$ dei polinomi con coefficienti in $\mathbb{Z}$ con 0 come *elemento neutro*, $-(a_0 + a_1 x + \dots + a_n x^n) := -a_0 - a_1 x - \dots - a_m x^m$ come *elemento inverso* di $a_0 + a_1 x + \dots + a_m x^m$, *somma* definita come $(a_0 + a_1 x + \dots + a_m x^m) + (b_0 + b_1 x + \dots + b_m x^m + b_{m+1} x^{m+1} + \dots + b_n x^n) = (a_0 + b_0) + (a_1 + b_1)x + \dots + (a_m + b_m)x^m + b_{m+1} x^{m+1} + \dots + b_n x^n$ (per $n \geq m$) e *prodotto* definito come $(a_0 + a_1 x + \dots + a_m x^m) \cdot (b_0 + b_1 x + \dots + b_m x^m + b_{m+1} x^{m+1} + \dots + b_n x^n) = \sum_{i=0}^{m+n} c_i x^i$, dove $c_i := \sum_{h+k=i} a_h b_k$ (es.: $c_0 = a_0 b_0, c_1 = a_0 b_1 + a_1 b_0, c_2 = a_0 b_2 + a_1 b_1 + a_2 b_0$) (vedasi Definizione 3.6 per una trattazione più rigorosa).

Gli esempi dati sono anelli commutativi.

- **Campi:** un *campo* nel linguaggio $\mathscr{L} := \{\dot{0}, \dot{1}, \dot{-}, \dotplus, *\}$ è una sestupla $\mathscr{K} := \langle K, \dot{0}, \dot{1}, \dot{-}, \dotplus, * \rangle$ con $\langle K, \dot{0}, \dot{-}, \dotplus, * \rangle$ anello commutativo tale che:

[3]Per semplificare la notazione stiamo supponendo che $*$ leghi più fortemente di $\dotplus$, analogamente al caso dell'algebra dei numeri reali, e quindi non abbiamo bisogno di scrivere $(x * y) \dotplus (x * z)$. Inoltre, sempre in analogia con il caso dei numeri reali, nella trattazione degli anelli potremo anche omettere la scrittura esplicita del simbolo del prodotto.

(k1) $(\forall x)\ \dot{1} * x = x = x * \dot{1}$ (elemento neutro)

(k2) $(\forall x)\,(x \neq \dot{0} \leftrightarrow (\exists y)\ x * y = \dot{1} = y * x)$ (elementi inversi)

Essendo quindi $\langle K \setminus \{\dot{0}\}, \dot{1}, *\rangle$ un gruppo, abbiamo che per ogni $x$ il suo elemento inverso rispetto a $*$ è univocamente determinato. Analogamente ai casi precedenti, questo fatto ci suggerisce di definire la nozione di campo anche in un linguaggio esteso contenente un simbolo $^{-1}$ per l'operazione di inversione rispetto a $*$ (forzando ad essere, diciamo, $\dot{0}^{-1} = \dot{0}$):

Un campo nel linguaggio $\mathscr{L} := \langle \dot{0}, \dot{1}, \dot{-}, ^{-1}, \dot{+}, *\rangle$ è una settupla $\mathscr{K} := \langle K, \dot{0}, \dot{1}, \dot{-}, ^{-1}, \dot{+}, *\rangle$ con $\langle K, \dot{0}, \dot{-}, \dot{+}, *\rangle$ anello commutativo che soddisfa l'assioma (k1) e:

(k2') $(\forall x)\,(x \neq \dot{0} \leftrightarrow x * x^{-1} = \dot{1} = x^{-1} * x)$ (elemento inverso)

*Esempi:*

1. $\langle \mathbb{Q}, 0, 1, -, ^{-1}, +, \cdot\rangle$

Nel seguito ci riserveremo di non distinguere notazionalmente tra una struttura e l'insieme su cui è definita nel caso delle strutture classiche studiate in matematica, come $\mathbb{Z}, \mathbb{Z}_p, \mathbb{Q}, \mathbb{R}, \mathbb{C}$, secondo il costume in uso in letteratura.

### *3.1.1 Anelli*

**Proposizione 3.1.** *Sia* $\mathscr{R} := \langle R, \dot{0}, \dot{-}, \dot{+}, *\rangle$ *un anello. Allora per tutti gli* $x, y, a \in R$ *vale*

1. $\dot{-}(x \dot{+} y) = \dot{-}x \dot{-} y$;
2. $x * \dot{0} = \dot{0} = \dot{0} * x$;
3. $\dot{-}(a * x) = (\dot{-}a) * x$.

*Dimostrazione.*

1. Esercizio (segue dall'associatività e dalla commutatività di $\dot{+}$ e dal fatto che gli elementi inversi sono univocamente determinati; il risultato vale quindi per gli anelli in quanto essi sono gruppi abeliani);

2. $x * y = x * (y \dot{+} 0) = (x * y) \dot{+} (x * \dot{0})$ per distributività a sinistra. Ma sappiamo anche che $(x * y) \dot{+} \dot{0} = x * y$. Per quanto osservato prima in merito ai gruppi (e cioè che l'elemento neutro, anche relativamente ai singoli individui, è univocamente determinato), concludiamo che $\dot{0} = x * \dot{0}$ (la dimostrazione di $\dot{0} = \dot{0} * x$ è analoga);

3. Per la distributività a destra vale $(a * x) \dot{+} [(\dot{-} a) * x] = (a \dot{-} a) * x = \dot{0} * x = \dot{0}$. Giacchè in un gruppo l'elemento inverso è sempre univocamente determinato per ciascun elemento, dev'essere $\dot{-}(a * x) = (\dot{-} a) * x$.

□

**Definizione 3.2** (Anello Unitario)**.** *Un* anello unitario *nel linguaggio* $\mathscr{L} := \{\dot{0}, \dot{1}, \dot{-}, \dot{+}, *\}$ *è una sestupla* $\mathscr{R} := \langle R, \dot{0}, \dot{1}, \dot{-}, \dot{+}, *\rangle$ *tale che* $\langle R, \dot{0}, \dot{-}, \dot{+}, *\rangle$ *è un anello ed inoltre l'assioma (k1) è soddisfatto*[4].

In virtù dell'assioma (k1), in un anello unitario contenente qualche $x \neq \dot{0}$ si deve avere $\dot{1} \neq \dot{0}$ per la Proposizione 3.1.2.

Inoltre in qualsiasi anello unitario la condizione $\dot{0} \neq \dot{1}$ è equivalente alla condizione $(\forall x)((\exists y) x * y = \dot{1} = y * x \rightarrow x \neq \dot{0})$ (esercizio; suggerimento, si usi la Proposizione 3.1.2).

### *3.1.2 Campi*

**Definizione 3.3** (Domini di Integrità)**.** *Un* dominio di integrità *è un anello commutativo unitario* $\mathscr{D} := \langle D, \dot{0}, \dot{1}, \dot{-}, \dot{+}, *\rangle$ *senza divisori dello zero, ovvero per ogni* $x, y \in D$, $x * y = \dot{0} \Rightarrow x = \dot{0}$ *o* $y = \dot{0}$.

**Fatto 3.4.** *Ogni campo è un dominio di integrità.*

*Dimostrazione.* Siano $x, y$ tali che $x * y = \dot{0}$. Supponiamo che $x \neq \dot{0}$. Allora esiste l'elemento inverso $x^{-1}$ di $x$ rispetto a $*$ per (k2'). Da questo segue $y = \dot{1} * y = x^{-1} * x * y = x^{-1} * \dot{0} = \dot{0}$ per la Proposizione 3.1.2. Il caso $y \neq \dot{0}$ è analogo. □

[4] Sebbene la maggior parte degli anelli considerati in letteratura siano in effetti unitari, esistono anche degli esempi importanti di anelli non unitari, ad esempio l'insieme delle funzioni a supporto compatto in $\mathbb{R}$.

Se pertanto $k$ non è primo, $\mathbb{Z}_k$ non può essere un campo perché ha divisori dello 0, esistono cioè $x, y \in \mathbb{Z}_k$ tali che $x \cdot y \mod k = 0$ sebbene $x \neq 0 \neq y$: esistono infatti $m, n$ per $1 < m, n < k$ tali che $k = m \cdot n$.

Invece per ogni $p$ primo, $\mathbb{Z}_p$ è il più piccolo campo di caratteristica $p$ a meno di isomorfismi, ovvero $\mathbb{Z}_p$ si immerge[5] in ogni altro campo di caratteristica $p$ (tale cioè per cui $\underbrace{\dot{1} \dotplus \dot{1} \dotplus \ldots \dotplus \dot{1}}_{p\text{-volte}} = \dot{0}$).

$\mathbb{Q}$ ha invece caratteristica 0, in quanto $\underbrace{1+1+\ldots+1}_{k\text{-volte}} \neq 0$ per ogni $k \geq 1$; non solo, è il più piccolo campo di caratteristica 0, nel senso che si immerge in ogni altro campo di caratteristica 0 (ponendo $\frac{m}{n} = m \cdot n^{-1}$).

**Esercizio 3.5.** *Trova un'interpretazione logica per le operazioni di somma e prodotto nel campo $\mathbb{Z}_2$.*

### *3.1.3 Campi algebricamente chiusi*

Per introdurre il concetto di campo algebrico chiuso ricordiamo succintamente la nozione di polinomio, rimandando il lettore ai testi di Algebra di base per una trattazione più rigorosa.

**Definizione 3.6** (Polinomi)**.** *Sia $\mathscr{K}$ un campo definito sull'insieme K (secondo uno dei due linguaggi considerati).*

*Un polinomio in $\mathscr{K}$ è una successione infinita $(k_0, k_1, k_2, \ldots)$ di elementi di K tale che per qualche $i \in \mathbb{N}$ abbiamo $k_j = 0$ per $j \geq i$.*

*L'insieme dei polinomi su $\mathscr{K}$ costituisce un anello unitario con*[6]

$$\dot{0} := (\dot{0}, \dot{0}, \dot{0}, \ldots)$$

$$\dot{1} = (\dot{1}, \dot{1}, \dot{1}, \ldots)$$

$$\dot{-}(k_0, k_1, k_2, \ldots) = (\dot{-}k_0, \dot{-}k_1, \dot{-}k_2, \ldots)$$

[5]Rimandiamo al Capitolo 5 per una definizione rigorosa del concetto di "immersione".

[6]Utilizziamo lo stesso linguaggio di $\mathscr{K}$, ma ovviamente l'interpretazione dei simboli del linguaggio cambia nelle due strutture. Come abbiamo stabilito, anche qui non distinguiamo notazionalmente tra i simboli del linguaggio e le loro interpretazioni.

$$(k_0, k_1, k_2, ...) \dot{+} (k'_0, k'_1, k'_2, ...) = (k_0 \dot{+} k'_0, k_1 \dot{+} k'_1, k_2 \dot{+} k'_2, ...)$$

$$(k_0, k_1, k_2, ...) * (k'_0, k'_1, k'_2, ...) = (k_0 k'_0, k_0 k'_1 + k_1 k'_0, ..., \sum_{h+l=i} k_h k'_l, ...)$$

*Possiamo rappresentare il polinomio* $(k_0, k_1, k_2, ...)$ *mediante la somma formale infinita* $A(x) := k_0 x^0 \dot{+} k_1 x^1 \dot{+} k_2 x^2 \dot{+} ...$ *in una variabile* $x$*, dove* $x^i$ *per* $i \geq 0$ *è il prodotto formale* $\underbrace{x * ... * x}_{i\text{-volte}}$ *e* $k_i x^i$ *è abbreviazione per il prodotto formale* $k_i * x^i$*. Coincidendo* $x^1$ *con* $x$ *e* $x^0$ *con la stringa di lunghezza 0, abbiamo che* $A(x) = k_0 \dot{+} k_1 x \dot{+} k_2 x^2 \dot{+} ....$

*Per ogni* $i \in \mathbb{N}$, $k_i$ *sarà il* coefficiente *dell'addendo* $k_i x^i$.

*Il polinomio* $(\dot{0}, \dot{0}, \dot{0}, ...)$ *è detto* polinomio nullo.

*Indichiamo con* $\mathscr{K}[x]$ *l'anello dei polinomi rappresentato mediante somme nella variabile* $x$.

*Per semplicità possiamo anche supporre che nelle somme* $A(x)$ *non debbano necessariamente occorrere (tutti) gli addendi con coefficienti nulli, ovvero del tipo* $\dot{0}\, x^i$*; in particolare possiamo ridurci a somme* $A(x)$ *finite (come caso limite, il polinomio nullo può essere rappresentato anche dalla somma vuota di lunghezza 0).*

*Nel seguito, identificheremo spesso il polinomio* $(k_0, k_1, k_2, ...)$ *con una qualsiasi somma* $A(x)$ *che lo rappresenti, spesso finita, come sarà chiaro dal contesto.*

*Il massimo* $m \leq n$ *tale che* $k_m \neq \dot{0}$ *(se esiste) è detto essere il* grado *del polinomio* $A(x)$ *(se tale* $m$ *non esiste, ovvero* $k_m = \dot{0}$ *per* $m \leq n$*, allora* $A(x)$ *è il polinomio nullo).*

Con somma (risp. prodotto) formale intendiamo un'espressione consistente in una somma di addendi (rispettivamente prodotto di fattori) che non può/non deve essere risolta. In particolare, la variabile $x$ non è un elemento del campo ma è un mero simbolo sintattico, quindi non ci attendiamo nessun risultato. Tuttavia questa variabile potrà essere interpretata sugli oggetti di $K$. In questo modo il polinomio $A(x) := k_0 \dot{+} k_1 x \dot{+} k_2 x^2 \dot{+} ...$ definisce una *funzione polinomiale* che per ogni $k \in K$ restituisce l'elemento $A(k) \in K$ tale che $A(k) = k_0 \dot{+} k_1 k \dot{+} k_2 k^2 \dot{+} ...$ per $k^i = \underbrace{k * ... * k}_{i\text{-volte}}$ per $i \geq 1$ e $k^0 = \dot{1}$. Si osservi pertanto che $A(k) = k_0 k^0 \dot{+} k_1 k^1 \dot{+} k_2 k^2 \dot{+} ....$

Si noti che $\dot{0}\, k = \dot{0}$ per ogni $k \in K$ per la Proposizione 3.1.2. Giacché $\dot{0}$ è elemento neutro di $\mathscr{K}$ rispetto a $\dot{+}$, la funzione polinomiale calcolata da $A(x)$ mediante interpretazione di $x$ sugli elementi del campo non cambia anche supponendo che essa non contenga alcuni (potenzialmente tutti) gli addendi con coefficienti nulli, quindi il concetto di funzione polinomiale è ben definito in ogni modo per $A(x)$.

Notazionalmente, potremo abbreviare $\dot{1}\, x^i$ come $x^i$. Per $k \in K$, con $kA(x)$ denoteremo il polinomio in cui i coefficienti di $A(x)$ sono stati tutti moltiplicati per $k$.

**Definizione 3.7** (Zeri (Radici) e Campi Algebricamente Chiusi)**.** *Sia $\mathscr{K}$ un campo definito su un insieme $K$. Allora $k \in K$ è detto essere uno* zero *(o* radice*) del polinomio $A(x)$ se $A(k) = \dot{0}$.*

*$\mathscr{K}$ è* algebricamente chiuso *se ogni polinomio di grado $m > 0$ ha una radice, o equivalentemente se ogni*

$$A(x) := k_n x^n \dot{+} k_{n-1} x^{n-1} \dot{+} \dots \dot{+} k_1 x + k_0$$

*con $k_n \neq 0$ per $n > 0$ possiede una radice.*

**Esempio 3.8.** *$\mathbb{C}$ è un campo algebricamente chiuso sulla base del Teorema Fondamentale dell'Algebra.*

### *3.1.4 Divisione di polinomi*

Descriviamo ora il metodo per la divisione di due polinomi in $\mathscr{K}[x]$ per un certo campo $\mathscr{K}$. Sia il polinomio

$$A(x) := a_n x^n \dot{+} \dots \dot{+} a_1 x \dot{+} a_0$$

da dividere per il polinomio

$$B(x) := b_m x^m \dot{+} \dots \dot{+} b_1 x \dot{+} b_0$$

per $n \geq m$ e $b_m \neq \dot{0}$.

Si eseguano i seguenti passaggi[7]:

[7] Per $a, b \in K$ e $b \neq \dot{0}$ si intenda $\frac{a}{b} = a * b^{-1}$. Trasferendo formalmente le regole degli elementi di $K$ su $x$, avremo che $\frac{x^n}{x^m} = x^n * (x^m)^{-1}$ (sotto la condizione formale $x^m \neq \dot{0}$ o più semplicemente $x \neq \dot{0}$ per la Proposizione 3.1.2 e il Fatto 3.4). Questo poi corrisponderà a $x^{n-m}$, per $n \geq m$. Per $m = 0$, la cosa è evidente in quanto $\dot{1}^{-1} = \dot{1}$. Per $m \geq 1$, la cosa segue dal fatto che in un gruppo gli elementi inversi sono univo-

(1) si divide il primo membro di $A(x)$ per il primo membro di $B(x)$ ottenendo il *quoziente provvisorio* $Q_1(x) := q_k x^k$:

$$\begin{array}{lcl} a_n x^n \dot{+} \ldots\ldots + a_1 x \dot{+} a_0 & | & b_m x^m \dot{+} \ldots\ldots \dot{+} b_1 x \dot{+} b_0 \\ & | & ------------ \\ & | & \underbrace{\frac{a_n}{b_m}}_{q_k} \underbrace{x^{n-m}}_{x^k} \end{array}$$

(2) si moltiplica $B(x)$ per $q_k x^k$ e si scrive il risultato $T(x)$ sotto $A(x)$[8]:

$$\begin{array}{lcl} a_n x^n \quad \dot{+} \ldots \dot{+} a_k x^k \dot{+} \ldots \dot{+} a_1 x \dot{+} a_0 & | & b_m x^m \dot{+} \ldots\ldots + b_1 x \dot{+} b_0 \\ & | & ------------ \\ b_m q_k \underbrace{x^{m+k}}_{x^n} \dot{+} \ldots \dot{+} b_0 q_k x^k & | & q_k x^k \end{array}$$

(3) si sottrae $T(x)$ da $A(x)$, ottenendo così il *resto provvisorio* $R_1(x)$:

$$\begin{array}{l} a_n x^n \quad \dot{+} a_{n-1} x^{n-1} \quad \dot{+} \ldots\ldots\ldots\ldots\ldots\ldots \dot{+} a_k x^k \dot{+} a_{k-1} x^{k-1} \dot{+} \ldots \dot{+} a_1 x \quad \dot{+} a_0 \\ \\ b_m q_k x^n \dot{+} b_{m-1} q_k x^{n-1} \dot{+} \ldots\ldots\ldots\ldots \dot{+} b_0 q_k x^k \\ ---------------------------------- \\ // \qquad r_{n-1} x^{n-1} \dot{+} \ldots\ldots\ldots\ldots\ldots\ldots \dot{+} r_k x^k \dot{+} \underbrace{r_{k-1}}_{a_{k-1}} x^{k-1} \dot{+} \ldots \dot{+} \underbrace{r_1}_{a_1} x \dot{+} \underbrace{r_0}_{a_0} \end{array}$$

A questo punto se $n-1 \geq m$ si ripetono i passaggi (1)–(3), con $R_1(x) = r_{n-1}x^{n-1} \dot{+} \ldots \dot{+} r_0$ al posto di $A(x)$, e $q_{k-1}x^{k-1} := \frac{r_{n-1}}{b_m} x^{n-1-m}$ al posto di $q_k x^k$, e $Q_2(x) = q_k x^k \dot{+} q_{k-1} x^{k-1}$ al posto di $Q_1(x)$.

Il processo si arresta quando si giunge ad un passo $N$ nel quale si arriva a scrivere come resto la somma vuota oppure si arriva ad un re-

camente determinati, e quindi $(x^m)^{-1} = (x^{-1})^m$ per (k1), (k2'), associatività e commutatività, in quanto $\underbrace{x * x * \ldots * x}_{m\text{-volte}} * \underbrace{x^{-1} * x^{-1} * \ldots * x^{-1}}_{m\text{-volte}} = \underbrace{(x * x^{-1}) * \ldots (x * x^{-1})}_{m\text{-volte}} = \dot{1}$.

Dunque $\frac{x^n}{x^m} = x^n * (x^m)^{-1} = x^n * (x^{-1})^m = x^{n-m}$ per associatività, commutatività, (k1) e (k2'). Evitiamo però di scrivere le espressioni $\frac{x^n}{x^m}$, $x^n * (x^m)^{-1}$ esplicitamente nei calcoli non essendo ammesse nella rappresentazione dei polinomi e scriviamo direttamente $x^{n-m}$.

[8] È immediato vedere che $x^{k+i} = x^k * x^i = x^{n-m} * x^i = x^{n-m+i}$ per $0 \leq i \leq m$ (per $i = m$ si ha quindi $x^{k+m} = x^n$).

sto che ha massima potenza della $x$ inferiore al grado $m$ di $B(x)$[9] (logicamente, dovrà quindi essere sempre il primo caso quando $m = 0$). Nel primo caso abbiamo in pratica che il resto $R_N(x)$ è il polinomio nullo.

Ogni qualvolta il resto coincide con il polinomio nullo[10], diciamo che $B(x)$ *divide* $A(x)$ e che $B(x)$ è un *divisore* di $A(x)$.

Si osservi che

$$R_1(x) = A(x) \dot{-} q_k x^k B(x)$$

e quindi gli zeri in comune a $B(x)$ e $A(x)$ sono anche zeri di $R_1(x)$ (si noti infatti che $q_k x^k B(x)$ viene annullato da tali zeri per Proposizione 3.1.2).

Viceversa, da

$$A(x) = q_k x^k B(x) \dot{+} R_1(x)$$

vediamo che gli zeri in comune a $B(x)$ e $R_1(x)$ sono anche zeri di $A(x)$.

Concludiamo che gli zeri in comune a $A(x)$ e $B(x)$ coincidono con quelli in comune a $R_1(x)$ e $B(x)$. Pertanto il passaggio nella divisione che porta da $A(x)$ e $B(x)$ a $R_1(x)$ e $B(x)$ preserva gli zeri.

Osserviamo tuttavia che, giacché $b_m \neq \dot{0}$, chiedersi se $k$ sia uno zero di $R_1(x)$ equivale a chiedersi se $k$ sia uno zero di $b_m R_1(x)$, dove

$$b_m R_1(x) = b_m A(x) \dot{-} a_n x^{n-m} B(x),$$

in quanto ogni campo $\mathscr{K}$ è un dominio di integrità per il Fatto 3.4. Pertanto gli zeri in comune a $A(x)$ e $B(x)$ sono quelli in comune a $b_m R_1(x)$ e $B(x)$. Useremo questo fatto nel commento al Teorema 7.4 dove utilizzeremo il linguaggio $\mathscr{L}$ per i campi senza $^{-1}$.

Vediamo ora come possiamo calcolare il massimo comun divisore tra $A(x)$ e $B(x)$. Un polinomio $M(x)$ è detto essere un *massimo comun divisore (m.c.d.)* di $A(x)$ e $B(x)$ se (i) divide sia $A(x)$ che $B(x)$ e (ii) qualsiasi polinomio $C(x)$ che divida $A(x)$ e $B(x)$ divide anche $M(x)$. Ebbene, partiamo dalla coppia $(A(x), B(x))$ ed avviamo il procedimento di divisione di $A(x)$ per $B(x)$, ottenendo il primo resto

---

[9] Stiamo ovviamente identificando il polinomio (infinito) del resto con la sua rappresentazione mediante la somma finita calcolata a quel passo: è propriamente questa ad avere un grado massimale della variabile $x$.

[10] Si noti che ciò si dà non solo per la somma vuota, ma anche ottenendo una somma con uno o più addendi tutti con coefficienti nulli.

provvisorio $R_1(x)$. Sostituiamo allora la prima coppia di polinomi con la nuova coppia $(R_1(x), B(x))$. Se la massima potenza di $x$ in $R_1(x)$ non è inferiore al grado di $B(x)$, si comincia a dividere $R_1(x)$ per $B(x)$, ottenendo un nuovo primo resto provvisorio $R_1'(x)$. Si prosegue in tale modo ottenendo ad ogni passo una nuova coppia, della quale un elemento sarà il primo resto provvisorio ottenuto a tale passo, mentre l'altro sarà ancora $B(x)$ (il polinomio per cui si divide). Ad ogni passo le massime potenze della $x$ dei resti ottenuti vanno diminuendo fino a che ad un certo punto otterremo un resto $R(x)$ che o sarà la somma vuota e quindi il polinomio nullo, oppure avrà una massima potenza della $x$ inferiore al grado di $B(x)$ (anche nel secondo caso, qualora tutti i coefficienti siano nulli, avremo il polinomio nullo!). Nel secondo caso, qualora $R(x)$ non sia nullo, possiamo dividere $B(x)$ per $R(x)$. Procedendo in modo analogo otterremo delle coppie costituite dai nuovi primi resti provvisori e dallo stesso $R(x)$ (il polinomio per cui si divide ora). Nuovamente, si perverrà ad un certo punto ad un resto $R'(x)$ che o è dato dalla somma vuota oppure avrà massima potenza della $x$ più piccola del grado di $R(x)$. Nel secondo caso, se questo non è nullo, si può cominciare a dividere $R(x)$ per $R'(x)$ e così via. Ebbene, prima o poi si perverrà al polinomio nullo (non necessariamente rappresentato dalla somma vuota). L'altro polinomio $M(x)$ che si troverà nella coppia corrispondente (ovvero il polinomio per il quale si dividerà in quel momento) sarà un massimo comun divisore di $A(x)$ e $B(x)$.

Per quanto osservato sulla preservazione delle radici, gli zeri di $A(x)$ e $B(x)$ sono gli zeri di $M(x)$ (nell'ultima coppia di polinomi è sufficiente considerare $M(x)$ in quanto ogni $k \in K$ è banalmente uno zero del polinomio nullo!).

Pertanto se $M(x)$ non ha radici, ad esempio se è il polinomio $k$ (cioè $kx^0$) per $k \neq \dot{0} \in K$, allora $A(x)$ e $B(x)$ non possono avere radici in comune. Altrimenti, se $M(x)$ ha radici in $\mathscr{K}$, allora anche $A(x)$ e $B(x)$ avranno radici in comune.

**Esempio 3.9.** *Si calcoli un massimo comun divisore delle seguenti coppie di polinomi in* $\mathbb{Q}[y]$ *con il metodo sopra illustrato:*

1. $2y, 4y$
2. $27y^3 - 75y, 3y^2 + 8y + 5 = 0$
3. $4y^2 + 2y$ *e* $5y + 10$

SOLUZIONI:

1. $$\begin{array}{l|l} 2y & 4y \\ 2y & ---\\ --- & \frac{1}{2} \\ // & \end{array}$$

Coppie di polinomi calcolate: $(2y,4y),(0,4y)$
m.c.d.: $4y$
(con 0 intendiamo qui il polinomio nullo $(0,0,0,\dots)$)

2. $$\begin{array}{l|l} 27y^3+0y^2-75y+0 & 3y^2+8y+5 \\ 27y^3+72y^2+45y+0 & --------- \\ ---------- & 9y \\ //-72y^2-120y+0 & \end{array}$$

$$\begin{array}{l|l} -72y^2-120y+0 & 3y^2+8y+5 \\ -72y^2-192y-120 & --------- \\ ---------- & -24 \\ //////72y+120 & \end{array}$$

$$\begin{array}{l|l} 3y^2+8y+5 & 72y+120 \\ 3y^2+5y+0 & --------- \\ -------- & \frac{1}{24}y \\ //////3y+5 & \end{array}$$

$$\begin{array}{l|l} 3y+5 & 72y+120 \\ 3y+5 & --------- \\ -------- & \frac{1}{24} \\ // & \end{array}$$

Coppie di polinomi calcolate: $(27y^3-75y,3y^2+8y+5)$, $(-72y^2-120y,3y^2+8y+5)$, $(72y+120,3y^2+8y+5)$, $(3y+5,72y+120)$, $(0,72y+120)$
m.c.d.: $72y+120$

3. $$\begin{array}{l|l} 4y^2+2y+0 & 5y+10 \\ 4y^2+8y+0 & --------- \\ -------- & \frac{4}{5}y \\ //-6y+0 & \end{array}$$

$$
\begin{array}{l|l}
-6y+0 & 5y+10 \\
-6y-12 & -------- \\
-------- & -\frac{6}{5} \\
//////12 &
\end{array}
$$

$$
\begin{array}{l|l}
5y+10 & 12 \\
5y & -------- \\
-------- & \frac{5}{12}y \\
////10 &
\end{array}
$$

$$
\begin{array}{l|l}
10 & 12 \\
10 & -------- \\
-------- & \frac{5}{6} \\
// &
\end{array}
$$

Coppie di polinomi calcolate: $(4y^2+2y, 5y+10)$, $(-6y, 5y+10)$, $(12, 5y+10)$, $(10,12)$, $(0,12)$
mcd: 12
(vediamo qui 12 come il polinomio $12y^0$)

### 3.1.5 *Spazi Vettoriali*

**Definizione 3.10** (Spazio Vettoriale)**.** *Uno* spazio vettoriale *su un campo* $\mathscr{K} := \langle K, 0, 1, -, ^{-1}, +, \times \rangle$[11] *è un gruppo abeliano* $\mathscr{V} := \langle V, \dot{0}, \dot{+} \rangle$ *associato ad un operazione di* prodotto scalare $\cdot : K \times V \to V$*, la quale soddisfa le seguenti proprietà (per semplicità notazionale, in accordo con la letteratura, possiamo omettere di scrivere in modo esplicito le operazioni di prodotto sul campo* $\times$ *e prodotto scalare* $\cdot$*):*

*(v1)* $1v = v$ *per ogni* $v \in V$

*(v2)* $k_1(k_2 v) = (k_1 k_2) v$ *per ogni* $k_1, k_2 \in K$ *e* $v \in V$

*(v3)* $k(v \dot{+} w) = kv \dot{+} kw$ *per ogni* $k \in K$ *e* $v, w \in V$

*(v4)* $(k_1 + k_2) v = k_1 v \dot{+} k_2 v$ *per ogni* $k_1, k_2 \in K$ *e* $v \in V$

[11] Usiamo in questo contesto per semplicità notazionale gli stessi simboli usati nell'algebra dei reali, il che non significa che il campo sia necessariamente $\mathbb{R}$ o un suo sottoinsieme.

*In (v1) l'elemento* $1$ *è da intendersi come l'elemento neutro nel campo rispetto al prodotto* $\times$ *(secondo l'assioma (k1)).*

Si osservi che la definizione di spazio vettoriale non è stata formulata in un linguaggio al primordine bensì in linguaggio informale. Questo perché il concetto di spazio vettoriale non è definibile nei nostri linguaggi al primordine: i quantificatori utilizzati nella sua definizione distinguono tra elementi di $K$ ed elementi di $V$ (la nozione di prodotto scalare è infatti definita su due strutture, non su una), distinzione inesprimibile mediante i nostri generici quantificatori $(\forall x)$, $(\exists x)$ che devono agire su un'unica struttura. Le condizioni (v1)-(v4) non vanno quindi considerate come assiomi al primordine, com'era invece per monoidi, gruppi, anelli e campi.

Anche se abbiamo dato la definizione di spazio vettoriale per gruppi nel solo linguaggio $\mathscr{L} := \{\dot{0}, \dot{+}\}$ in quanto questo linguaggio è sufficiente alla definizione, nel seguito utilizzeremo come di consueto il simbolo $\dot{-}$ per denotare l'operazione di inversione rispetto a $\dot{+}$, sulla base del fatto che come sappiamo gli elementi inversi in un gruppo comunque esistono e sono univocamente determinati per ogni singolo elemento.

**Proposizione 3.11.** *Sia* $\mathscr{V}$ *uno spazio vettoriale con dominio* $V$ *sul campo* $\mathscr{K}$ *con dominio* $K$*. Allora*

1. $0v = \dot{0} = k\dot{0}$ *per ogni* $v \in V, k \in K$*;*
2. $\dot{-}\, v = (-1)v$ *e* $\dot{-}\,(kv) = (-k)v$ *per ogni* $v \in V, k \in K$*;*
3. $kv = \dot{0} \Rightarrow k = 0$ *oppure* $v = \dot{0}$ *per ogni* $v \in V, k \in K$*.*

*Dimostrazione.*

1. Per (v4) abbiamo $0v = (0+0)v = 0v \dot{+} 0v$. Per unicità dell'elemento neutro nel gruppo $\mathscr{V}$ abbiamo che $0v = \dot{0}$. La seconda equazione si dimostra in modo analogo (usando però (v3)) o si deriva dalla prima mediante (v2): $k\dot{0} = k(0v) = (k0)v = 0v = \dot{0}$ per ogni $v \in V$ (infatti $\mathscr{K}$ è un campo, per cui $k0 = 0$ per Proposizione 3.1.2));

2. Per (v4) e per il punto 1 abbiamo $kv \dot{+} (-k)v = (k-k)v = 0v = \dot{0}$; per unicità degli elementi inversi nel gruppo $\mathscr{V}$ concludiamo

quindi $(-k)v = \dot{-}(kv)$. La dimostrazione di $\dot{-}v = (-1)v$ si ottiene come caso particolare per $k := 1$ previa osservazione che $v = 1v$ per (v1);

3. Supponiamo che $kv = \dot{0}$ ma $k \neq 0$. Allora per (v1), (v2) ed il punto 1 abbiamo $v = 1v = (k^{-1}k)v = k^{-1}(kv) = k^{-1}\dot{0} = \dot{0}$.

□

Fondamentale nello studio degli spazi vettoriali è il concetto di base:

**Definizione 3.12** (Basi)**.** *Sia dato uno spazio vettoriale $\mathscr{V}$ con dominio $V$ su un campo $\mathscr{K}$ con dominio $K$. Una* combinazione lineare *di elementi di $V$ è una somma del tipo $k_1v_1 \dot{+} \ldots \dot{+} k_nv_n$, per qualche $n \in \mathbb{N}$, $k_1, \ldots, k_n \in K$, $v_1, \ldots, v_n \in V$.*

*Il vettore $v := k_1v_1 \dot{+} \ldots \dot{+} k_nv_n$ è detto allora essere* combinazione lineare *di $v_1, \ldots, v_n$.*

*Un insieme di vettori $\{v_i\}_{i\in I} \subseteq V$ è* linearmente indipendente *se nessun $v_i$ per $i \in I$ è combinazione lineare di vettori in $\{v_j\}_{j\in I\setminus\{i\}}$.*

*Una* base *per $\mathscr{V}$ è un insieme linearmente indipendente $B \subseteq V$ che genera l'intero $V$, tale cioè che per ogni $v \in V$ esistono $v_1, \ldots, v_n \in B$ e $k_1, \ldots, k_n \in K$, per un qualche $n \in \mathbb{N}$, tali che $v = k_1v_1 \dot{+} \ldots \dot{+} k_nv_n$.*

**Osservazione 3.13.** *Una combinazione lineare si può sempre trasformare in* forma normale*, ovvero in una combinazione lineare del tipo $k_1v_1 \dot{+} \ldots \dot{+} k_nv_n$ per $v_1, \ldots, v_n$ distinti, mediante raccoglimenti (assioma (v4)).*

Il seguente teorema rappresenta una delle più celebri applicazioni dell'Assioma di Scelta in Matematica (nella forma del Lemma di Zorn):

**Teorema 3.14.** *Ogni spazio vettoriale possiede una base.*

Si osservi che dato uno spazio vettoriale $\mathscr{V}$ su un campo $\mathscr{K}$ con base $B$, ed un qualsiasi vettore $v$, esiste una sola combinazione lineare (a meno di commutazioni!) di elementi di $B$ *in forma normale e a coefficienti non nulli* (cioè tutti i coefficienti in essa contenuti sono diversi da 0) che dia come risultato $v$.

Per prima cosa, osserviamo che se

$$a_1v_1 \dot{+} \ldots \dot{+} a_nv_n = \dot{0},$$

per $a_1 v_1 \dotplus \ldots \dotplus a_n v_n$ in forma normale e $v_1, \ldots, v_n \in B$, allora $a_i = 0$ per $1 \leq i \leq n$. Supponiamo infatti che ciò non sia; sia ad esempio $a_1 \neq 0$. Allora abbiamo che $a_1 v_1 = \dot{-}\, a_2 v_2 \,\dot{-}\, \ldots \,\dot{-}\, a_n v = (-a_2) v_2 \dotplus \ldots \dotplus (-a_n) v_n$, da cui $v_1 = \frac{-a_2}{a_1} v_2 \dotplus \ldots \dotplus \frac{-a_n}{a_1} v_n$. Dunque $B$ non sarebbe un insieme linearmente indipendente.

Pertanto, se $v \neq \dot{0}$, una qualsiasi combinazione lineare in forma normale di vettori della base che lo esprima deve contenere qualche coefficiente non nullo, ed in particolare almeno una di queste deve avere tutti i coefficienti non nulli (basta cancellare gli addendi con coefficiente nullo sulla base del fatto che $0v' = \dot{0}$ per ogni $v' \in V$). Dimostriamo ora che ve n'è esattamente una di tal tipo. Sia infatti per assurdo $v \neq \dot{0}$ esprimibile mediante due combinazioni lineari normali distinte a coefficienti non nulli di elementi di $B$, diciamo $v = a_1 v_1 \dotplus \ldots \dotplus a_m v_m = b_{m+1} v_{m+1} \dotplus \ldots \dotplus b_n v_n$ con $a_i \neq 0 \neq b_j$ per $1 \leq i \leq m, m+1 \leq j \leq n$. Per uniformità, estendiamo se necessario le due combinazioni lineari in modo che entrambe siano combinazioni lineari normali definite sugli stessi vettori $v_1, \ldots, v_m, v_{m+1}, \ldots, v_n$ aggiungendo addendi con coefficienti nulli dove occorre. Otteniamo così $v = a_1 v_1 \dotplus \ldots \dotplus a_n v_n = b_1 v_1 \dotplus \ldots \dotplus b_n v_n$. Poiché le due combinazioni lineari iniziali normali differivano ma non contenevano coefficienti nulli, è facile vedere che vi è almeno un vettore $v_i$ con due coefficienti diversi nelle due nuove combinazioni lineari (uno dei quali, ma non entrambi, può ora essere nullo). Sia ad esempio, diciamo, $a_1 \neq b_1$. Allora $\dot{0} = v \,\dot{-}\, v = a_1 v_1 \dotplus \ldots \dotplus a_n v_n \,\dot{-}\, b_1 v_1 \,\dot{-}\, \ldots \,\dot{-}\, b_n v_n = (a_1 v_1 \,\dot{-}\, b_1 v_1) \dotplus (a_2 v_2 \dotplus \ldots \dotplus a_n v_n \,\dot{-}\, b_2 v_2 \,\dot{-}\, \ldots \,\dot{-}\, b_n v_n)$. Perciò $a_1 v_1 \,\dot{-}\, b_1 v_1 = a_1 v_1 \dotplus (-b_1) v_1 = (a_1 - b_1) v_1 = \dot{-}\, a_2 v_2 \,\dot{-}\, \ldots \,\dot{-}\, a_n v_n \dotplus b_2 v_2 \dotplus \ldots \dotplus b_n v_n = (-a_2) v_2 \dotplus \ldots \dotplus (-a_n) v_n \dotplus b_2 v_2 \dotplus \ldots \dotplus b_n v_n = (b_2 - a_2) v_2 \dotplus \ldots \dotplus (b_n - a_n) v_n$, da cui $v_1 = \frac{(b_2 - a_2)}{a_1 - b_1} v_2 \dotplus \ldots \dotplus \frac{(b_n - a_n)}{a_1 - b_1} v_n$ (si ricordi che $a_1 \neq b_1 \Rightarrow a_1 - b_1 \neq 0$). Pertanto $B$ non sarebbe in realtà linearmente indipendente (esercizio: mostrare la validità delle uguaglianze utilizzate in queste argomentazioni unicamente sulla base della definizione di spazio vettoriale e dei risultati precedentemente dimostrati).

La proprietà si può estendere anche al vettore $\dot{0}$ associandogli il caso limite di una combinazione lineare di vettori di $B$ di lunghezza 0: tutti i coefficienti contenuti in essa sono, banalmente, non nulli (giacché non ve n'è alcuno di nullo!) e tutti i vettori in essa contenuti sono distinti (giacché non ce ne sono due uguali!): dobbiamo ricorrere a ciò, perché abbiamo visto come qualsiasi combinazione lineare in forma normale di elementi della base che denoti $\dot{0}$ e che contenga qualche addendo debba avere tutti i coefficienti nulli.

Il seguente teorema ci consente l'introduzione del concetto di dimensione di uno spazio vettoriale:

**Teorema 3.15.** *Due basi qualsiasi di uno spazio vettoriale hanno la stessa cardinalità.*

**Definizione 3.16** (Dimensione)**.** *Sia $k$ un cardinale. Uno spazio vettoriale $\mathscr{V}$ ha* dimensione $k$ *($\dim(\mathscr{V}) = k$) se $\mathscr{V}$ ha una base di cardinalità $k$.*

**Esempio 3.17.**

1. *$\mathbb{R}$ è un gruppo abeliano che può essere visto come uno spazio vettoriale su sé stesso, in quanto campo. Come tale $\mathbb{R}$ ha dimensione 1, in quanto generato dal vettore $1$ (o da qualsiasi altro numero reale).*

2. *Dato un campo $\mathscr{K}$, l'insieme $\mathscr{K}^n$ di tutte le $n$-uple di elementi del dominio $K$ di $\mathscr{K}$ può essere visto come spazio vettoriale sul campo $\mathscr{K}$ avente dimensione $n$. Ha infatti come base standard l'insieme dei vettori $v_1, \ldots, v_n$ dove $v_i$ è l'$n$-upla di elementi di $K$ il cui $i$-esimo elemento è 1 e tutti gli altri sono 0, per $1 \leq i \leq n$. Le operazioni di somma vettoriale e prodotto scalare sono definite nel seguente modo:*

$$v := \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, w = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \mapsto v \dot{+} w := \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix}$$

$$k,\ v := \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \mapsto kv := \begin{pmatrix} ka_1 \\ \vdots \\ ka_n \end{pmatrix}$$

3. *$\mathbb{R}$ può anche essere visto come spazio vettoriale sul campo $\mathbb{Q}$, e in tal caso ha dimensione $2^{\aleph_0}$. Infatti $\mathbb{Q}$ è numerabile, e se consideriamo un qualsiasi sottoinsieme numerabile $B \subseteq \mathbb{R}$, l'insieme di tutte le combinazioni lineari in forma normale e a coefficienti non nulli definibile tramite $B$ è un sottoinsieme dell'insieme di tutte le stringhe finite sull'alfabeto $\mathbb{Q} \cup B \cup \{+\}$, e come tale*

*può essere solamente numerabile. Pertanto, poiché $\mathbb{R}$ contiene $2^{\aleph_0}$ elementi, per generare $2^{\aleph_0}$ combinazioni lineari (di tale tipo) dobbiamo avere una base $B \subseteq \mathbb{R}$ con altrettanti elementi.*

**Definizione 3.18** (Spazi Vettoriali Isomorfi)**.** *Due spazi vettoriali $\mathscr{V}_1$, $\mathscr{V}_2$, con domini $V_1$ e $V_2$ rispettivamente, e definiti su uno stesso campo $\mathscr{K}$ con dominio $K$ sono detti* isomorfi *se esiste una funzione biettiva $h : V_1 \to V_2$ che sia* lineare*, ovvero*[12]*:*

*(i)* $h(v \dot{+} w) = h(v) \dot{+} h(w)$ *per ogni* $v, w \in V_1$*;*

*(ii)* $h(k \cdot v) = k \cdot h(v)$ *per ogni* $v \in V_1, k \in K$.

**Proposizione 3.19.** *Siano dati due spazi vettoriali isomorfi $\mathscr{V}_1, \mathscr{V}_2$ con domini $V_1$ e $V_2$ rispettivamente e definiti sul campo $\mathscr{K}$ con dominio $K$. Per $h : V_1 \to V_2$ isomorfismo tra $\mathscr{V}_1$ e $\mathscr{V}_2$ si dà:*

1. $h(\dot{-}\, v) = \dot{-}\, h(v)$;
2. $h(\dot{0}) = \dot{0}$.

*Dimostrazione.*

1. Per la Proposizione 3.11.2 $h(\dot{-}v) = h((-1)v) = (-1)h(v) = \dot{-}h(v)$;
2. Esercizio (utilizzare la Proposizione 3.11.1).

□

**Teorema 3.20** (della Dimensione)**.** *Due spazi vettoriali $\mathscr{V}_1, \mathscr{V}_2$ su uno stesso campo $\mathscr{K}$ sono isomorfi sse hanno la stessa dimensione.*

*Dimostrazione.* $\Rightarrow$) Supponiamo che $\mathscr{V}_1$ e $\mathscr{V}_2$ spazi vettoriali su $\mathscr{K}$, con domini $V_1$ e $V_2$ rispettivamente, siano isomorfi mediante l'isomorfismo $h : V_1 \to V_2$. Sia $B_1$ una base di $\mathscr{V}_1$. È immediato vedere che $h(B_1)$ è una base di $\mathscr{V}_2$ e che ha la stessa cardinalità di $B_1$ (esercizio).

$\Leftarrow$) Siano $V_1$ e $V_2$ i domini di $\mathscr{V}_1$ e $\mathscr{V}_2$ rispettivamente, e siano date due basi $B_1 := \{v_i\}_{i\in I}$, $B_2 := \{v'_i\}_{i\in I}$ di $\mathscr{V}_1$ e $\mathscr{V}_2$ rispettivamente (per $I$

[12] Senza perdita di generalità assumiamo che $\mathscr{V}_1$ e $\mathscr{V}_2$ in quanto gruppi siano definiti sullo stesso linguaggio. Quest'assunzione semplifica la notazione ed aumenta l'analogia con il caso degli isomorfisimi che definiremo nel Capitolo 5 per le strutture definibili al primordine.

non necessariamente numerabile). La funzione $h : B_1 \to B_2, v_i \mapsto v'_i$ per $i \in I$ è una biezione (si osservi che per come abbiamo definito il concetto di insieme linearmente indipendente si ha $v_j \neq v_k, v'_j \neq v'_k$ per $j \neq k$). Estendiamo tale $h$ ad una funzione su tutto $V_1$ nel seguente modo: sia $v := k_1 v_{i_1} \dot{+} \ldots \dot{+} k_n v_{i_n}$ per $i_1, \ldots, i_n \in I$ e $k_1 v_{i_1} \dot{+} \ldots \dot{+} k_n v_{i_n}$ combinazione lineare di elementi della base $B_1$ in forma normale a coefficienti non nulli (cioè $k_1, \ldots, k_n \neq 0$); allora poniamo

$$h(v) := k_1 v'_{i_1} \dot{+} \ldots \dot{+} k_n v'_{i_n}.$$

(per quanto riguarda $h(\dot{0})$, ricordiamo la convenzione di considerare la combinazione lineare di lunghezza 0 come unico modo possibile di rappresentare il vettore $\dot{0}$; questo ci darà $h(\dot{0}) = \dot{0}$ ponendo nella definizione di sopra $n = 0$). La funzione è ben definita, perché già sappiamo che ogni vettore $v$ è esprimibile mediante una sola combinazione $k_1 v_{i_1} \dot{+} \ldots \dot{+} k_n v_{i_n}$ del tipo suddetto (considerando anche la possibilità di $n = 0$)[13].

Inoltre $h$ è iniettiva, sempre per il fatto che ciascun vettore (in $V_1, V_2$) è esprimibile mediante una sola combinazione lineare in forma normale di elementi di $B_1$ a coefficienti non nulli, e differenti combinazioni lineari di tale tipo vengono banalmente trasformate, per definizione di $h$, in combinazioni lineari del medesimo tipo a loro volta differenti.

Il completamento della dimostrazione che $h$ è un isomorfismo è lasciata come esercizio. □

In particolare, ciascuno spazio vettoriale $\mathscr{V}$ di dimensione $n$ finita definito su un campo $\mathscr{K}$ è isomorfo a $\mathscr{K}^n$. Per vedere questo è sufficiente identificare ciascun vettore $v := k_1 v_1 \dot{+} \ldots \dot{+} k_n v_n$, per $\{v_1, .., v_n\}$

[13] Abbiamo dato la definizione di $h(v)$ utilizzando solo combinazioni lineari a coefficienti non nulli unicamente per semplificare la dimostrazione del fatto che $h$ è ben definita e iniettiva. Ma al fine dell'ottenimento del valore desiderato, nulla cambierebbe se aggiungessimo a $k_1 v_{i_1} \dot{+} \ldots \dot{+} k_n v_{i_n}$ degli addendi con coefficienti nulli. Poiché infatti $0 v_j = \dot{0}$ per $v_j \in B_1$, l'aggiunta di tali coefficienti non cambierebbe la denotazione della combinazione lineare estesa, che denoterebbe sempre $v$, né questo altererebbe la definizione di $h(v)$, in quanto ciascun addendo $0 v_j$ *sarebbe* trasformato, per la nuova definizione conseguente di $h$, nell'addendo $0 v'_j$ per $v'_j \in B_2$, e, nuovamente, $0 v'_j = \dot{0}$. Come caso particolare, questa definizione confermerebbe la scelta di $h(\dot{0}) = \dot{0}$.

base selezionata per $\mathscr{V}$, con il vettore $h(v)$ di $\mathscr{K}^n$:

$$h(v) = \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix}$$

Si può vedere facilmente che la funzione $h$ così definita è un isomorfismo tra $\mathscr{V}$ e $\mathscr{K}^n$, cioè è biettiva ed inoltre $h(v \dotplus w) = h(v) \dotplus h(w)$ e $h(kv) = kh(v)$.

## 3.2 Spazi topologici

**Definizione 3.21** (Spazio topologico)**.** *Uno* spazio topologico *è una coppia* $(X, \tau_X)$ *dove* $X$ *è un insieme e* $\tau_X \subseteq \wp(X)$ *è una* topologia *su* $X$*, ovvero una collezione di sottoinsiemi di* $X$ *tale che:*

1. $\tau_X$ *è chiusa per unioni arbitrarie, ovvero se* $O_i \in \tau_X$ *per ogni* $i \in I$*, allora* $\bigcup_{i \in I} O_i \in \tau_X$

2. $\tau_X$ *è chiusa per intersezioni finite, ovvero dato* $I$ *finito se* $O_i \in \tau_X$ *per ogni* $i \in I$*, allora* $\bigcap_{i \in I} O_i \in \tau_X$

*Gli elementi di* $\tau_X$ *sono detti* insiemi aperti.

*Un sottoinsieme* $A \subseteq X$ *è detto* chiuso *se è il complementare di qualche insieme aperto.*

*Un sottoinsieme* $B \subseteq X$ *sia aperto che chiuso è detto* clopen.

*Per semplicità, identificheremo spesso* $X$ *con* $(X, \tau_X)$.

**Osservazione 3.22.** *Dalla condizione 1 segue che* $\emptyset \in \tau_X$.

*Dalla condizione 2 segue che* $X \in \tau_X$.

*Pertanto gli insiemi* $\emptyset$ *e* $X$ *sono clopen.*

**Esempio 3.23.** *Dato* $X$*, la topologia più povera su* $X$ *è data da* $\{\emptyset, X\}$ *(*topologia banale*).*

**Definizione 3.24** (Basi)**.** *Una* base *per uno spazio topologico* $(X, \tau_X)$ *è data da una collezione* $\{O_i\}_{i \in I}$ *di elementi di* $\tau_X$ *tali che per ogni* $O \in \tau_X$ *esiste* $J \subseteq I$ *tale che* $O = \bigcup_{i \in J} O_i$.

*Una base per* $\tau_X$ *è detta* generare $\tau_X$.

**Esempio 3.25.** *1. Dato $X$, la topologia più ricca su $X$ è quella generata da $\{\{x\} \mid x \in X\}$; questa è detta* topologia discreta *e coincide con $\wp(X)$.*

*2. La topologia ordinaria sui numeri reali viene generata dalla collezione di tutti gli intervalli aperti in $\mathbb{R}$, ovvero la collezione di tutti gli insiemi del tipo $]a,b[:= \{c \in \mathbb{R} \mid a < c < b\}$ per $a < b \in \mathbb{R}$.*

*Gli intervalli chiusi $[a,b] := \{c \in \mathbb{R} \mid a \leq c \leq b\}$ per $a < b \in \mathbb{R}$ sono tutti insiemi chiusi rispetto a tale topologia.*

**Definizione 3.26.** *Dato un punto $x \in X$ ed un sottoinsieme $B$ di $X$, abbiamo che:*

- *$x$ è punto* interiore *di $B$ se esiste un aperto $O \subseteq X$ tale che $x \in O \subseteq B$;*
- *$x$ è punto di* chiusura *di $B$ se per ogni aperto $O \subseteq X$ tale che $x \in O$ si ha $O \cap B \neq \emptyset$;*
- *$x$ è punto* limite *di $B$ se per ogni aperto $O \subseteq X$ tale che $x \in O$ si ha $(O \setminus \{x\}) \cap B \neq \emptyset$.*

Senza perdita di generalità possiamo supporre che nella Definizione 3.26 $O$ sia un aperto della base (selezionata) di $\tau_X$. Infatti da un lato ogni aperto di base è un aperto; dall'altra, se $x \in O$ allora esiste un aperto di base $O'$ tale che $x \in O' \subseteq O$.

Si noti che ogni punto limite di un insieme è anche un suo punto di chiusura, mentre non vale il viceversa. Si noti anche che ogni punto di un insieme è un suo punto di chiusura.

**Notazione 3.27.** *Dato un insieme $B \subseteq X$ indichiamo con*

- $B^\circ := \left\{x \in X \mid x \text{ è punto interiore di } B\right\}$
- $\overline{B} := \left\{x \in X \mid x \text{ è punto di chiusura di } B\right\}$
- $B' := \left\{x \in X \mid x \text{ è punto limite di } B\right\}$

**Fatto 3.28.** *Per ogni $B \subseteq X$ si ha:*

1. *$B \subseteq \overline{B}$ (ogni punto di un insieme è un suo punto di chiusura)*

2. *$B' \subseteq \overline{B}$ (ogni punto limite di un insieme è anche un suo punto di chiusura)*

*Dimostrazione.* Segue immediatamente dalla Definizione 3.26. □

**Lemma 3.29.** *1. Un insieme $O \subseteq X$ è aperto sse contiene esattamente i suoi punti interiori (ovvero $O = O^\circ$).*

*2. Un insieme $A \subseteq X$ è chiuso sse contiene esattamente i suoi punti di chiusura (ovvero $A = \overline{A}$).*

*Dimostrazione.*

1. Esercizio.

2. $\Rightarrow$) Da un lato abbiamo $A \subseteq \overline{A}$ per Fatto 3.28.1, quindi ogni punto di $A$ è un suo punto di chiusura. Vediamo ora che $x \in \overline{A} \Longrightarrow x \in A$, ovvero ogni punto di chiusura di $A$ è in $A$. Supponiamo infatti che $x \notin A$. Allora $x \in O := (X \setminus A)$ e $O$ è aperto. Ma $A \cap O = \emptyset$, e quindi $x$ non è un punto di chiusura di $A$, cioè $x \notin \overline{A}$.

   $\Leftarrow$) Sia $x \in (X \setminus A)$ (se tale $x$ non esiste, allora $A = X$ è chiuso e siamo a posto). Giacché $x \notin A$, per ipotesi $x$ non può essere un punto di chiusura di $A$, pertanto esiste un aperto $O$ tale che $x \in O$ e $O \cap A = \emptyset$. Pertanto $x \in O \subseteq (X \setminus A)$. Ne concludiamo che $X \setminus A$ è un unione di insiemi aperti (inclusi in esso), pertanto è aperto. Da cui $A$ è chiuso.

□

**Definizione 3.30** (Compattezza)**.** *Un insieme $K \subseteq X$ è detto* compatto *se ogni ricoprimento aperto di $K$ ha un sottoricoprimento finito, ovvero se per ogni famiglia $\{O_i\}_{i\in I}$ di insiemi aperti tale che $K \subseteq \bigcup_{i\in I} O_i$ esistono $O_{i_1}, \dots, O_{i_n} \in \{O_i\}_{i\in I}$ per un qualche $n \in \mathbb{N}$ tali che $K \subseteq \bigcup_{1\le j\le n} O_{i_j}$.*

*Qualora $K = X$ diciamo che $K$ è uno* spazio topologico compatto.

La seguente proposizione fornisce una carattterizzazione degli spazi topologici compatti[14]. A tal scopo diciamo che una famiglia di in-

[14] Si potrebbe generalizzare questo risultato a un generico sottoinsieme compatto $K \subseteq X$ per un $X$ spazio topologico arbitrario mediante la *topologia relativa* indotta da $X$ su $K$ (ovvero la topologia con aperti gli insiemi $O \cap K$ tali che $O$ è aperto di $X$).

siemi $\{B_i\}_{i\in I}$ ha la *proprietà dell'intersezione finita (p.i.f.)* se ogni sottoinsieme finito di tale famiglia ha intersezione non vuota, ovvero se per ogni $B_{i_1},\dots,B_{i_n} \in \{B_i\}_{i\in I}$ si ha $\bigcap_{1\le j\le n} B_{i_j} \ne \emptyset$.

**Proposizione 3.31.** *Uno spazio topologico $(K,\tau_K)$ è compatto sse ogni famiglia $\{A_i\}_{i\in I}$ di sottoinsiemi chiusi di K con la p.i.f. ha intersezione non vuota (ovvero $\bigcap_{i\in I} A_i \ne \emptyset$).*

*Dimostrazione.* $\Rightarrow$) Supponiamo che vi sia una famiglia $\{A_i\}_{i\in I}$ di sottoinsiemi chiusi di $K$ con la p.i.f. con intersezione vuota. Allora $\{K \setminus A_i\}_{i\in I}$ è un ricoprimento aperto di $K$ che non contiene nessun ricoprimento finito, quindi $K$ non è compatto.

$\Leftarrow$) Supponiamo che $K$ non sia compatto. Esiste quindi un ricoprimento aperto $\{O_i\}_{i\in I}$ di $K$ che non contiene nessun ricoprimento finito. Allora $\{K \setminus O_i\}_{i\in I}$ è una famiglia di insiemi chiusi di $K$ con la p.i.f. la cui intersezione è vuota. □

**Definizione 3.32** (Continuità)**.** *Una funzione $f : X \to Y$ tra spazi topologici è* continua in $x \in X$ *se per ogni aperto $U \subseteq Y$ tale che $f(x) \in U$ esiste un aperto $O \subseteq X$ con $x \in O$ tale che $f(O) \subseteq U$.*

*$f$ è continua (in $X$) se è continua in ogni $x \in X$.*

Intuitivamente, $f$ è continua in $x \in X$ se volendo "approssimare" $f(x)$ con un grado di "precisione arbitraria", possiamo arrivare all'obiettivo "approssimando sufficientemente bene" $x$. Anche in questo caso possiamo ridurci a considerare solo aperti delle basi (selezionate) di $X$ ed $Y$. Per $X = Y = \mathbb{R}$ possiamo quindi scrivere la condizione di continuità di $f$ in $x$ nel seguente modo: $f$ è continua in $x \in \mathbb{R}$ se per ogni $\varepsilon > 0$ esiste un $\delta > 0$ tale che per ogni $y \in \mathbb{R}$: $|x-y| \le \delta \Rightarrow |f(x)-f(y)| \le \varepsilon$. Coerentemente, $f$ è continua (in $\mathbb{R}$) se tale condizione vale per ogni $x \in \mathbb{R}$, ovvero se per ogni $x \in X$ e per ogni $\varepsilon > 0$ esiste un $\delta > 0$ tale che per ogni $y \in \mathbb{R}$: $|x-y| \le \delta \Rightarrow |f(x)-f(y)| \le \varepsilon$.

Il concetto di continuità non va confuso con quello di *continuità uniforme*, che ne costituisce un caso particolare proprio: $f$ è *uniformemente continua* (in $\mathbb{R}$) se per ogni $\varepsilon > 0$ esiste un $\delta > 0$ tale che per ogni $x, y \in \mathbb{R}$: $|x-y| \le \delta \Rightarrow |f(x)-f(y)| \le \varepsilon$. Intuitivamente, nel caso della continuità, il grado di approssimazione richiesto per l'argomento $x$ (cioè il $\delta$) viene detto dipendere da $x$ stesso preso singolarmente, mentre nel caso della continuità uniforme viene detto esserci uno stesso grado di approssimazione $\delta$ che va bene per tutti gli argomenti $x$ (ai fini dell'approssimazione dei valori della funzione a meno di $\varepsilon$)[15].

Possiamo generalizzare le corrispondenti nozioni anche al caso di funzioni definite su domini $D \subseteq \mathbb{R}$: in tal caso consideriamo solamente punti $x, y \in D$[16].

[15] Al riguardo, si noti il diverso ordine dei quantificatori "per ogni" ed "esiste" riferiti a $x$ e $\delta$ rispettivamente nelle due definizioni.

[16] Rigorosamente, questo consiste nel restringere le definizioni date alla *topologia relativa* indotta da $\mathbb{R}$ su $D$, ovvero la topologia generata dagli insiemi $]a, b[ \cap D$ per $a, b \in \mathbb{R}$ e $a < b$.

# Capitolo 4

# Teorema di Compattezza

## *4.1 Compattezza logica*

**Definizione 4.1** (Soddisfacibile e Finitamente Soddisfacibile)**.** *Un insieme di $\mathscr{L}$-formule chiuse* $\Gamma$ *è detto* soddisfacibile *se possiede un modello, ovvero esiste una struttura* $\mathscr{M}$ *tale che* $\mathscr{M} \models \varphi$ *per ogni* $\varphi \in \Gamma$.

*Un insieme di $\mathscr{L}$-formule chiuse* $\Gamma$ *è detto* finitamente soddisfacibile *se ogni suo sottoinsieme finito è soddisfacibile.*

Sebbene il Teorema di Compattezza valga per linguaggi $\mathscr{L}$ di cardinalità arbitraria, per semplicità lo dimostreremo, per ora, solo per linguaggi $\mathscr{L}$ numerabili (finiti o infiniti). Il caso generale, per linguaggi di cardinalità infinita arbitraria, può essere dimostrato generalizzando la dimostrazione che diamo qui sotto estendendo il principio di induzione anche al caso degli ordinali limiti. Pertanto tutti gli ingredienti necessari per provare il caso generale sono essenzialmente già contenuti nella dimostrazione qui esposta. Tuttavia, noi daremo in seguito una dimostrazione per il caso generale di natura differente (Teorema 9.11).

**Teorema 4.2** (Teorema di Compattezza (caso numerabile))**.** *Sia $\mathscr{L}$ un linguaggio numerabile. Un insieme* $\Gamma$ *di $\mathscr{L}$-formule chiuse è soddisfacibile se e solo se è finitamente soddisfacibile.*

Dimostreremo il teorema in due parti. Nella prima considereremo solamente formule che non contengano il simbolo di uguaglianza

"=". Nella seconda parte estenderemo il risultato a formule che possano contenerlo. In entrambi i casi la direzione *solo se* (⇒) è banale. Occorre pertanto soffermarsi solamente sulla direzione *se* (⇐).

*Caso senza* =:

*Dimostrazione.* Sia dato un insieme $\Gamma$ di $\mathscr{L}$-formule senza "=" finitamente soddisfacibile. L'idea di base della dimostrazione è quella di costruire un insieme $\overline{\Gamma} \supseteq \Gamma$ che contenga in sé sufficienti informazioni utili per costruire un modello dell'intero $\Gamma$.

A tal fine aggiungiamo al linguaggio $\mathscr{L}$ una quantità infinita numerabile di nuove costanti distinte $c_0, c_1, \ldots, c_n, \ldots$. Otteniamo così un nuovo linguaggio $\mathscr{L}'$ che pur estendendo $\mathscr{L}$ è pur sempre numerabile.

Sia ora $\varphi_0, \varphi_1, \varphi_2, \ldots$ un'enumerazione di tutte le $\mathscr{L}'$-formule chiuse senza =.

Definiamo ora per induzione una successione di insiemi di $\mathscr{L}'$-formule $\Gamma_0, \Gamma_1, \Gamma_2, \ldots$ tali che $\Gamma_0 = \Gamma$, $\Gamma_i \subseteq \Gamma_{i+1}$, $\Gamma_i$ è finitamente soddisfacibile per ogni $i \in \mathbb{N}$:

- $\Gamma_0 := \Gamma$
- $$\Gamma_{i+1} := \begin{cases} \Gamma_i \cup \{\varphi_i\} & \text{se } \Gamma_i \cup \{\varphi_i\} \text{ è fin. soddisfacibile} \\ & \text{e } \varphi_i \text{ non è della forma } (\exists x)\psi \\ \\ \Gamma_i \cup \{\varphi_i, \psi(c/x)\} & \text{se } \Gamma_i \cup \{\varphi_i\} \text{ è fin. soddisfacibile,} \\ & \varphi_i \text{ è della forma } (\exists x)\psi \\ & \text{e } c \text{ è la costante di indice minimo} \\ & \text{in } \{c_0, c_1, \ldots, c_n, \ldots\} \\ & \text{a non essere contenuta in } \Gamma_i \cup \{\varphi_i\} \\ \\ \Gamma_i \cup \{\neg\varphi_i\} & \text{se } \Gamma_i \cup \{\varphi_i\} \text{ non è fin. sodd.} \end{cases}$$

Si noti che la definizione di $\Gamma_{i+1}$ è consistente perchè $\Gamma_0$ non contiene alcuna costante in $\{c_0, c_1, \ldots, c_n, \ldots\}$ e ad ogni passo induttivo $j < i+1$ al massimo una nuova costante viene aggiunta a $\Gamma_j$, per cui al passo $i+1$ abbiamo ancora infinite costanti nuove a disposizione.

Infine poniamo $\overline{\Gamma} := \bigcup_{i \in \mathbb{N}} \Gamma_i$.

Vediamo che:

- $\Gamma \subseteq \overline{\Gamma}$: ovvio, perché $\Gamma = \Gamma_0 \subseteq \overline{\Gamma}$;

- $\Gamma_i$ è finitamente soddisfacibile per ogni $i \in \mathbb{N}$:

  - $i = 0$: per ipotesi

  - $i = j + 1$: se $\Gamma_j \cup \{\varphi_j\}$ è finitamente soddisfacibile e $\varphi_j$ non è della forma $(\exists x)\psi$, allora $\Gamma_i := \Gamma_j \cup \{\varphi_j\}$, pertanto la proprietà è già dimostrata. Sia invece $\Gamma_j \cup \{\varphi_j\}$ finitamente soddisfacibile e $\varphi_j \equiv (\exists x)\psi$. Allora $\Gamma_i := \Gamma_j \cup \{\varphi_j, \psi(c/x)\}$. Prendiamo $\Gamma' \subseteq \Gamma_i$ finito. Senza perdita di generalità possiamo supporre che $\{\varphi_j, \psi(c/x)\} \subseteq \Gamma'$. Poniamo $\Gamma'' := \Gamma' \setminus \{\psi(c/x)\}$. Poiché $\Gamma'' \subseteq \Gamma_j \cup \{\varphi_j\}$ e $\Gamma_j \cup \{\varphi_j\}$ è finitamente soddisfacibile, esiste un modello $\mathscr{M}$ di $\Gamma''$, che possiamo considerare definito rispetto al linguaggio (minimale) $\mathscr{L}^*$ di $\Gamma''$. In particolare $\mathscr{M} \models \psi(x)[m]$ per qualche $m \in M$, giacché $(\exists x)\psi \in \Gamma''$. Estendiamo $\mathscr{M}$ ad una struttura $\mathscr{M}'$ sul linguaggio $\mathscr{L}^* \cup \{c\}$ mediante l'interpretazione $c^{\mathscr{M}'} := m$. Allora banalmente $\mathscr{M}' \models \Gamma''$ (in quanto in $\Gamma''$ non occorre $c$!) ed inoltre $\mathscr{M}' \models \psi(c/x)$ (dimostrazione per induzione: esercizio), da cui $\mathscr{M}' \models \Gamma'$. Sia infine $\Gamma_j \cup \{\varphi_j\}$ non finitamente soddisfacibile. Allora esiste un sottoinsieme finito $\Gamma' \subseteq \Gamma_j$ tale che $\Gamma' \cup \{\varphi_j\}$ non è soddisfacibile. Per definizione $\Gamma_i := \Gamma_j \cup \{\neg\varphi_j\}$. Supponiamo per assurdo che anche $\Gamma_j \cup \{\neg\varphi_j\}$ non sia finitamente soddisfacibile. Allora esiste anche un sottoinsieme finito $\Gamma'' \subseteq \Gamma_j$ tale che $\Gamma'' \cup \{\neg\varphi_j\}$ non è soddisfacibile. Poiché $\Gamma' \cup \Gamma'' \subseteq \Gamma_j$ e $\Gamma_j$ è finitamente soddisfacibile per ipotesi di induzione, esiste una struttura $\mathscr{M}$ tale che $\mathscr{M} \models \Gamma' \cup \Gamma''$. Se necessario, sia $\mathscr{M}'$ un'estensione banale arbitraria di $\mathscr{M}$ compatibile con il linguaggio di $\varphi_j$, altrimenti $\mathscr{M}' = \mathscr{M}$. Ovviamente si dà $\mathscr{M}' \models \varphi_j$ oppure $\mathscr{M}' \models \neg\varphi_j$ (in quanto $\varphi_j$ è chiusa). Nel primo caso $\mathscr{M}' \models \Gamma' \cup \{\varphi_j\}$, il che contraddice la nostra ipotesi. Nel secondo caso $\mathscr{M}' \models \Gamma'' \cup \{\neg\varphi_j\}$, e anche questo contraddice la nostra ipotesi. Pertanto $\Gamma_j \cup \{\neg\varphi_j\}$ dev'essere finitamente soddisfacibile;

- $\overline{\Gamma}$ è finitamente soddisfacibile: supponiamo per assurdo che non lo sia. Allora esiste un sottoinsieme finito $\{\psi_1, \ldots, \psi_n\} \subseteq \overline{\Gamma}$

non soddisfacibile. Giacchè $\overline{\Gamma} = \bigcup_{i\in\mathbb{N}} \Gamma_i$ e $\Gamma_0 \subseteq \Gamma_1 \subseteq \Gamma_2 \subseteq .... \subseteq \Gamma_i \subseteq \Gamma_{i+1} \subseteq ...$ esiste un $k$ tale che $\{\psi_1,...,\psi_n\} \subseteq \Gamma_k$ (dimostrazione per esercizio). Per il punto precedente sappiamo che $\Gamma_k$ è finitamente soddisfacibile, perciò $\{\psi_1,...,\psi_n\}$ dev'essere soddisfacibile. Contraddizione;

- $\overline{\Gamma}$ è un *insieme di Henkin*, cioè:
    - $\psi \in \overline{\Gamma}$ oppure $\neg\psi \in \overline{\Gamma}$ per ogni $\mathscr{L}'$-formula chiusa $\psi$: data $\psi$, questa coinciderà con la formula $\varphi_k$ per un certo $k$. Allora $\varphi_k \in \Gamma_{k+1} \subseteq \overline{\Gamma}$ oppure $\neg\varphi_k \in \Gamma_{k+1} \subseteq \overline{\Gamma}$;
    - se $(\exists x)\psi \in \overline{\Gamma}$ allora esiste un termine chiuso $t$ tale che $\psi(t/x) \in \overline{\Gamma}$: sia infatti $(\exists x)\psi$ la formula $\varphi_k$ per un certo $k$. Consideriamo cosa succede al passo $k+1$. Se $\Gamma_k \cup \{\varphi_k\}$ non è finitamente soddisfacibile, allora $\Gamma_{k+1} := \Gamma_k \cup \{\neg\varphi_k\}$, e quindi $\overline{\Gamma}$ conterrebbe l'insieme finito insoddisfacibile $\{\varphi_k, \neg\varphi_k\}$, il che contraddice quanto sopra dimostrato. Pertanto $\Gamma_k \cup \{\varphi_k\}$ dev'essere finitamente soddisfacibile, e quindi $\Gamma_{k+1} := \Gamma_k \cup \{\varphi_k, \psi(c/x)\}$, dove $c$ è la prima delle costanti in $c_0, c_1, ..., c_n, ...$ non contenuta in $\Gamma_k \cup \{\varphi_k\}$.

Vediamo ora che possiamo costruire un modello $\mathscr{M}$ di $\overline{\Gamma}$ definito su un insieme $M$ fatto di *oggetti linguistici*. Poniamo:

- $M$ è l'insieme dei termini chiusi di $\mathscr{L}'$[1];
- $c^{\mathscr{M}} := c$ per ogni costante $c$ di $\mathscr{L}'$;
- $(f^n)^{\mathscr{M}} : M^n \to M, (t_1,..,t_n) \mapsto f^n(t_1,...,t_n)$,
  cioè $(f^n)^{\mathscr{M}}(t_1,..,t_n) := f^n(t_1,..,t_n)$;
- $(t_1,...,t_n) \in (P^n)^{\mathscr{M}} \subseteq M^n$ sse $P^n(t_1,..,t_n) \in \overline{\Gamma}$.

Dimostriamo ora per induzione che per ogni $\mathscr{L}'$-formula chiusa $\varphi$ si dà $\mathscr{M} \models \varphi$ se e solo se $\varphi \in \overline{\Gamma}$:

[1] Si osservi che per questo caso particolare potremo quindi scrivere semplicemente $\mathscr{M} \models \varphi(t_1,...,t_n)$ per ogni formula $\varphi$ anziché $\mathscr{M} \models \varphi(x_1,...,x_n)[t_1,..,t_n]$, in quanto $t_1,..,t_n$ oltre ad essere oggetti del dominio della stuttura saranno anche termini del linguaggio.

- $\varphi \equiv P(t_1,...,t_n)$: $\mathscr{M} \models P(t_1,...,t_n)$ sse $(t_1,...,t_n) \in P^{\mathscr{M}}$ sse $P(t_1,..,t_n) \in \overline{\Gamma}$ per la definizione di $\mathscr{M}$;

- $\varphi \equiv \bot$: $\mathscr{M} \not\models \bot$ e $\bot \notin \overline{\Gamma}$ (perché $\{\bot\}$ è insieme finito insoddisfacibile)

- $\varphi \equiv \neg\psi$: $\mathscr{M} \models \neg\psi$ sse $\mathscr{M} \not\models \psi$ sse $\psi \notin \overline{\Gamma}$ (per IH) sse $\neg\psi \in \overline{\Gamma}$ (essendo $\overline{\Gamma}$ un insieme di Henkin);

- $\varphi \equiv \psi \vee \xi$: $\mathscr{M} \models \psi \vee \xi$ sse $\mathscr{M} \models \psi$ o $\mathscr{M} \models \xi$ sse $\psi \in \overline{\Gamma}$ o $\xi \in \overline{\Gamma}$ (per IH). Vediamo che: $\psi \in \overline{\Gamma}$ o $\xi \in \overline{\Gamma}$ sse $\psi \vee \xi \in \overline{\Gamma}$:

  $\Rightarrow$) sia $\psi \vee \xi \notin \overline{\Gamma}$ allora $\neg(\psi \vee \xi) \in \overline{\Gamma}$ (essendo $\overline{\Gamma}$ di Henkin). Giacché $\{\psi, \neg(\psi \vee \xi)\}, \{\xi, \neg(\psi \vee \xi)\}$ sono insiemi finiti insoddisfacibili, allora sia $\psi \notin \overline{\Gamma}$ sia $\xi \notin \overline{\Gamma}$, altrimenti $\overline{\Gamma}$ non sarebbe finitamente soddisfacibile, il che contraddirebbe quanto prima dimostrato;

  $\Leftarrow$) sia $\psi, \xi \notin \overline{\Gamma}$. Allora $\neg\psi, \neg\xi \in \overline{\Gamma}$ (essendo $\overline{\Gamma}$ di Henkin). Giacché $\{\neg\psi, \neg\xi, \psi \vee \xi\}$ è un insieme finito insoddisfacibile, allora $\psi \vee \xi \notin \overline{\Gamma}$, altrimenti $\overline{\Gamma}$ non sarebbe finitamente soddisfacibile, il che contraddirebbe quanto prima dimostrato;

- $\varphi \equiv \psi \wedge \xi$: questo caso si riduce ai precedenti considerando che da un lato $\mathscr{M} \models \psi \wedge \xi \Leftrightarrow \mathscr{M} \models \neg(\neg\psi \vee \neg\xi)$ e dall'altro $\psi \wedge \xi \in \overline{\Gamma} \Leftrightarrow \neg(\neg\psi \vee \neg\xi) \in \overline{\Gamma}$ (essendo $\overline{\Gamma}$ di Henkin e finitamente soddisfacibile);

- $\varphi \equiv \psi \to \xi$: anche questo caso si riduce ai precedenti in modo analogo utilizzando l'equivalenza logica di $\psi \to \xi$ e $\neg\psi \vee \xi$;

- $\varphi \equiv (\exists x)\psi$: $\mathscr{M} \models (\exists x)\psi$ sse $\mathscr{M} \models \psi(t/x)$ per qualche $\mathscr{L}'$-termine chiuso $t$ sse $\psi(t/x) \in \overline{\Gamma}$ (per IH). Vediamo che $\psi(t/x) \in \overline{\Gamma}$ per qualche $\mathscr{L}'$-termine chiuso $t$ sse $(\exists x)\psi \in \overline{\Gamma}$:

  $\Rightarrow$) sia $(\exists x)\psi \notin \overline{\Gamma}$. Allora $\neg(\exists x)\psi \in \overline{\Gamma}$ (essendo $\overline{\Gamma}$ di Henkin). Giacché $\{\psi(t/x), \neg(\exists x)\psi\}$ è un insieme finito insoddisfacibile per ogni termine (chiuso) $t$, allora $\psi(t/x) \notin \overline{\Gamma}$ per ogni termine (chiuso) $t$, altrimenti $\overline{\Gamma}$ non sarebbe finitamente soddisfacibile, il che contraddirebbe quanto prima dimostrato;

  $\Leftarrow$) sia $(\exists x)\psi \in \overline{\Gamma}$. Allora esiste un termine chiuso $t$ tale che $\psi(t/x) \in \overline{\Gamma}$, in quanto $\overline{\Gamma}$ è insieme di Henkin;

- $\varphi \equiv (\forall x)\psi$: questo caso si riduce ai precedenti utilizzando l'equivalenza logica di $(\forall x)\psi$ e $\neg(\exists x)\neg\psi$.

Dunque $\mathscr{M} \models \overline{\Gamma}$, ed in particolare $\mathscr{M} \models \Gamma$. Per ottenere un modello di $\Gamma$ che sia una $\mathscr{L}$-struttura è sufficiente ripulire $\mathscr{M}$ dalle funzioni di interpretazione delle nuove costanti $c_0, c_1, ..., c_n, ....$ □

Come sappiamo, nella logica del primordine con identità il simbolo di uguaglianza = è una costante logica esprimente l'identità (sul dominio della struttura). Chiamiamo $AxId$ gli assiomi $(I1)-(I5)$ istanziati per tutti i simboli di funzioni e predicati di $\mathscr{L}$. Si noti banalmente che per ogni struttura $\mathscr{A}$ si dà $\mathscr{A} \models AxId$ (infatti, ad esempio, non esiste alcun oggetto nel dominio di $\mathscr{A}$ che possa essere diverso da sé stesso, e quindi la riflessività è necessariamente soddisfatta, e così via)[2]. Tuttavia alla base dell'interpretazione "forzata" del simbolo = come identità vi sta una convenzione, in quanto gli assiomi $(I1)-(I5)$ non caratterizzino la nozione di identità in modo inequivocabile (al contrario di quanto accade per gli altri simboli logici, connettivi e quantificatori, all'interno del sistema di assiomi), e se rimpiazziamo = con un nuovo simbolo, diciamo $\simeq$, tale convenzione verrà immediatamente a cadere. Data una $\mathscr{L}$-formula $\varphi$ denotiamo con $\varphi_\simeq$ la formula del linguaggio esteso $\mathscr{L}_\simeq$ ottenuta sostituendo in $\varphi$ il simbolo = con $\simeq$, ed analogamente per un qualsiasi insieme di $\mathscr{L}$-formule $\Gamma$. Possiamo tuttavia chiamare il nuovo sistema di assiomi $AxId_\simeq$ così ottenuto anche come $AxCongr$. La ragione per tale nome è che una qualsiasi relazione che soddisfi tali assiomi sarà una relazione di congruenza, senza che debba essere per forza l'identità, essendo venuta meno la convenzione suddetta. L'identità è in effetti una relazione di congruenza, ma l'inverso non vale (quindi l'indentità è solo un *caso particolare* di congruenza). Intuitivamente, due oggetti (in una struttura) sono identici se soddisfano esattamente le stesse proprietà tra tutte quelle possibili, e non solamente quelle denominate dal linguaggio della struttura (ovvero le sole proprietà di fatto esistenti nella struttura). Per definire l'identità ci vuole pertanto il linguaggio al second'ordine, mediante il quale possiamo definirla nel seguente modo: $x = y \equiv_{def} (\forall P)(P(x) \leftrightarrow P(y))$. Al primordine

[2] Si ricordi che il sistema di assiomi per la logica del primordine con identità che abbiamo presentato soddisfa il Teorema di Validità, pertanto tutti i suoi assiomi, tra cui $(I1)-(I_5)$, devono essere necessariamente validi (su tutte le strutture).

possiamo tutt'al più caratterizzare la congruenza mediante gli assiomi $AxCongr$, che è una nozione relativa al linguaggio usato. Come dire, *la maglia tessuta al primordine è troppo larga per calzare perfettamente sulla sola identità. Ciò che possiamo imbrigliare è la sola congruenza.* Due oggetti sono congruenti quando sono indistinguibili rispetto a tutte le proprietà e funzioni di cui parla il linguaggio. Come dicevamo, due oggetti identici sono per forza congruenti, ma il contrario non vale. Ad esempio, supponiamo di avere un linguaggio $\mathscr{L}$ contenente solamente un simbolo di predicato unario $P$ ed un simbolo di funzione unario $d$. Consideriamo una struttura il cui dominio sia $\mathbb{N}$ e dove $P$ venga interpretato come "pari" e $d$ come una funzione che restituisce il resto della divisione per due del suo argomento. Allora in tale $\mathscr{L}$-struttura tutti i numeri pari sono tra loro congruenti, cosiccome lo sono tra loro tutti i numeri dispari. Eppure due numeri pari distinti qualsiasi, come pure due dispari, non soddisfano le stesse proprietà tra tutte quelle possibili. Infatti due numeri pari (risp. dispari) $n$ ed $m$ tali che $n \neq m$ sono distinguibili mediante le due proprietà differenti $\{m\}, \{n\} \subseteq \mathbb{N}$. In generale, individui diversi saranno sempre distinguibili mediante i loro rispettivi singoletti, ma non è detto che quest'ultimi appartengano alla struttura.

Si noti ancora che sia l'identità che qualsiasi congruenza sono relazioni di equivalenza, cioè sono relazioni riflessive, simmetriche e transitive: l'identità soddisfa infatti $(I1), (I2), (I3)$, e una qualsiasi congruenza soddisfa gli assiomi $(I1)_{\simeq}, (I2)_{\simeq}, (I3)_{\simeq}$.

Consideriamo ora che $\Gamma$ sia un insieme di formule della logica con identità, intendendo dire con ciò che le formule di $\Gamma$ *possono* contenere il simbolo $=$. Supponiamo che alcune di queste formule siano del tipo $t_1 = t_2$ per $t_1$ e $t_2$ termini *sintatticamente distinti.* Se costruissimo analogamente al caso senza identità una struttura $\mathscr{M}$ mediante un insieme di Henkin $\overline{\Gamma}$ ottenuto a partire da $\Gamma$, $t_1$ e $t_2$ verrebbero interpretati nel dominio di tale struttura su sé stessi, e quindi $\mathscr{M} \not\models t_1 = t_2$. Pertanto tale struttura non potrebbe essere un modello di $\Gamma$!

Se però sostituiamo $=$ con $\simeq$ allora è sufficiente che $t_1$ e $t_2$ siano congruenti. Vediamo ora come utilizzare questo fatto per costruire una nuova struttura $\mathscr{M}'$ a partire da $\mathscr{M}$ che sia effettivamente un modello per $\Gamma$ con identità. Dimostreremo in questo modo la direzione *se* del caso con identità.

*Caso con =*

*Dimostrazione.* sia per ipotesi $\Gamma$ (possibilmente) contentente = finitamente soddisfacibile. Ciò significa che dato $\Gamma' \subseteq \Gamma$ finito, esiste $\mathscr{M}_{\Gamma'}$ tale che $\mathscr{M}_{\Gamma'} \models \Gamma'$. Sappiamo anche che $\mathscr{M}_{\Gamma'} \models \Gamma' \cup AxId$. A questo punto è possibile aggiungere a $\mathscr{M}_{\Gamma'}$ un'interpretazione di $\simeq$ in modo da ottenere un modello $\mathscr{M}_{\Gamma'_{\simeq}}$ di $\Gamma'_{\simeq} \cup AxCongr$ rispetto al linguaggio $\mathscr{L}_{\simeq} := \mathscr{L} \cup \{\simeq\}$. Ciò può essere banalmente fatto: è sufficiente aggiungere a $\mathscr{M}_{\Gamma'}$ in modo esplicito il predicato di identità $I := \{(m,m) : m \in M_{\Gamma'}\}$, per $M_{\Gamma'}$ dominio di $\mathscr{M}_{\Gamma'}$, ed intepretare $\simeq$ come $I$. Questo significa che $\Gamma_{\simeq} \cup AxCongr$ è finitamente soddisfacibile. Per la dimostrazione del Teorema di Compattezza nel caso senza = concludiamo che $\Gamma_{\simeq} \cup AxCongr$ è soddisfacibile. Sia quindi $\mathscr{M}$ un modello di $\Gamma_{\simeq} \cup AxCongr$.

Procediamo ora introducendo un metodo generale che poi applicheremo al nostro caso particolare: data una qualsiasi $\mathscr{L}_{\simeq}$-struttura $\mathscr{M}$ tale che $\mathscr{M} \models AxCongr$, diamo indicazioni su come costruire a partire da $\mathscr{M}$ una precisa $\mathscr{L}$-struttura $[\mathscr{M}]$ nel vecchio linguaggio $\mathscr{L}$.

Osserviamo che $\simeq^{\mathscr{M}}$ è una relazione di equivalenza sul dominio $M$, in quanto $\mathscr{M} \models \{(I1)_{\simeq}, (I2)_{\simeq}, (I3)_{\simeq}\}$, da cui si ottiene

1. $m_1 \simeq^{\mathscr{M}} m_1$ (riflessività),
2. $m_1 \simeq^{\mathscr{M}} m_2 \Rightarrow m_2 \simeq^{\mathscr{M}} m_1$ (simmetria),
3. $m_1 \simeq^{\mathscr{M}} m_2 \;\&\; m_2 \simeq^{\mathscr{M}} m_3 \Rightarrow m_1 \simeq^{\mathscr{M}} m_3$ (transitività)

per ogni $m_1, m_2, m_3 \in M$.

Possiamo allora definire la struttura $[\mathscr{M}]$ basata sull'*insieme quoziente* $M/\simeq^{\mathscr{M}} = \{[m] \mid m \in M\}$, dove $[m] = \left\{m' \in M \mid m' \simeq^{\mathscr{M}} m\right\}$ per ogni $m \in M$, nel seguente modo:

- $c^{[\mathscr{M}]} := [c^{\mathscr{M}}]$ per ogni $\mathscr{L}$-costante $c$;
- $f^{[\mathscr{M}]}([m_1], ..., [m_n]) := [f^{\mathscr{M}}(m_1, .., m_n)]$ per ogni $\mathscr{L}$-simbolo di funzione $f$ (la nozione è ben definita per l'assioma $(I4_{\simeq})$

  $(\forall x_1)...(\forall x_n)(\forall y_1)...(\forall y_n)(x_1 \simeq y_1 \wedge ... \wedge x_n \simeq y_n \to f(x_1, ..., x_n) \simeq f(y_1, ..., y_n)))$;

- $([m_1],...,[m_n]) \in P^{[\mathscr{M}]}$ sse $(m_1,..,m_n) \in P^{\mathscr{M}}$ per ogni $\mathscr{L}$-simbolo di predicato $P$ (la nozione è ben definita per l'assioma $(I5_{\simeq})$ $(\forall x_1)...(\forall x_n)(\forall y_1)...(\forall y_n)(x_1 \simeq y_1 \wedge ... \wedge x_n \simeq y_n \to (P(x_1,...,x_n) \to P(y_1,...,y_n)))$.

**Lemma 4.3.** *Sia data una $\mathscr{L}$-formula $\varphi(x_1,...,x_n)$ ed una $\mathscr{L}_{\simeq}$-struttura $\mathscr{M}$ su un insieme M tale che $\mathscr{M} \models AxCongr$. Per ogni $m_1,...,m_n \in M$ si ha $\mathscr{M} \models \varphi_{\simeq}[m_1,...,m_n]$ sse $[\mathscr{M}] \models \varphi[[m_1],...,[m_n]]$.*

*Dimostrazione.* Per prima cosa bisogna dimostrare per induzione sulla costruzione dei termini che per ogni $\mathscr{L}$-termine $t(x_1,...,x_n)$ e per ogni $m_1,...,m_n \in M$ si ha

$$t^{[\mathscr{M}]}([m_1],...,[m_n]) = [t^{\mathscr{M}}(m_1,...,m_n)] \quad (\star)$$

(esercizio). Analizziamo poi qualche caso significativo dell'induzione sulla costruzione delle formule, il resto è lasciato per esercizio:

- $\varphi(x_1,...,x_n) \equiv P(x_1,...,x_n)$, per $P$ $\mathscr{L}$-simbolo di predicato: segue dalla definizione di $[\mathscr{M}]$ (in questo caso ovviamente $\varphi_{\simeq} \equiv \varphi$)

- $\varphi(x_1,...,x_n) \equiv t_1(x_1,...,x_n) = t_2(x_1,...,x_n)$: allora

$$\mathscr{M} \models (t_1 \simeq t_2)[m_1,...,m_n]$$

  sse $t_1^{\mathscr{M}}(m_1,...,m_n) \simeq^{\mathscr{M}} t_2^{\mathscr{M}}(m_1,...,m_n)$ sse $[t_1^{\mathscr{M}}(m_1,...,m_n)] = [t_2^{\mathscr{M}}(m_1,...,m_n)]$ sse $t_1^{[\mathscr{M}]}([m_1],...,[m_n]) = t_2^{[\mathscr{M}]}([m_1],...,[m_n])$ (per $\star$) sse $[\mathscr{M}] \models t_1 = t_2[[m_1],...,[m_n]]$

- $\varphi(x_1,...,x_n) \equiv (\exists x)\psi(x,x_1,...,x_n)$: allora

$$\mathscr{M} \models (\exists x)\psi_{\simeq}(x,x_1,...,x_n)[m_1,..,m_n]$$

  sse $\mathscr{M} \models \psi_{\simeq}(x,x_1,...,x_n)[m,m_1,...,m_n]$ per qualche $m \in M$ sse $[\mathscr{M}] \models \psi(x,x_1,...,x_n)[[m],[m_1],...,[m_n]]$ per qualche $[m] \in M/\simeq^{\mathscr{M}}$ (per IH) sse $[\mathscr{M}] \models (\exists x)\psi(x,x_1,...,x_n)[[m_1],...,[m_n]]$.

□

Sia quindi, come abbiamo supposto sopra, $\mathscr{M} \models \Gamma_{\simeq} \cup AxCongr$. Allora $[\mathscr{M}] \models \Gamma$ per il Lemma 4.3. Pertanto $\Gamma$ è soddisfacibile. Questo completa la dimostrazione del Teorema di Compattezza anche per il caso dell'identità[3]. □

## 4.2 *Alcune importanti conseguenze del Teorema di Compattezza*

Il Teorema di Compattezza è uno dei teoremi fondamentali della logica del primordine. Esso ha molte applicazioni *fruttuose* in logica e matematica (ad esempio l'esistenza di modelli non standard per l'analisi reale contenenti numeri illimitati ed infinitesimi in atto). Noi ci occuperemo al momento di alcune sue conseguenze logiche "negative", inerenti limitazioni espressive del calcolo al primordine, che comunque testimoniano la non esistenza di certe proprietà matematiche.

**Teorema 4.4** (Teorema di Skolem-Löwenheim verso l'alto (caso numerabile))**.** *Sia* $\Gamma$ *un insieme di* $\mathscr{L}$*-formule chiuse della logica con identità, per* $\mathscr{L}$ *numerabile. Se* $\Gamma$ *possiede modelli finiti di cardinalità arbitrariamente grande, allora ha un modello infinito.*

*Dimostrazione.* Aggiungiamo ad $\mathscr{L}$ un insieme infinito numerabile di nuove costanti distinte $d_0, d_1, d_2, \ldots$, ottenendo il linguaggio $\mathscr{L}'$. Consideriamo ora l'insieme di $\mathscr{L}'$-formule

$$\hat{\Gamma} := \Gamma \cup \left\{ d_i \neq d_j \mid i, j \in \mathbb{N}, i \neq j \right\}.$$

Vediamo ora che $\hat{\Gamma}$ è finitamente soddisfacibile. Sia infatti $\Gamma' \subseteq \hat{\Gamma}$ finito. Il caso interessante è quando $\Gamma'$ contiene delle nuove costanti, siano queste $d_{i_1}, \ldots, d_{i_n}$. Per ipotesi $\Gamma$ ha un modello $\mathscr{M}$ che contiene almeno $n$ elementi distinti tra loro $m_1, .., m_n$. Sia $\mathscr{M}'$ ottenuto da $\mathscr{M}$ aggiungendo le funzioni di interpretazione per le costanti $d_{i_j}$ in modo che $d_{i_j}^{\mathscr{M}'} := m_j$ per $1 \leq j \leq n$. È immediato vedere che $\mathscr{M}' \models \Gamma'$. Concludiamo che $\hat{\Gamma}$ è finitamente soddisfacibile. Per il

[3] Si osservi che se $\Gamma$ non contiene $=$ allora $\Gamma_{\simeq}$ non contiene $\simeq$. In questo caso, giacché comunque $\mathscr{M} \models \Gamma_{\simeq} \cup AxCongr$, la struttura $\mathscr{M}$ deve contenere l'interpretazione di $\simeq$ e soddisfare $(I1)_{\simeq}$. Per queste ragioni si darà $[m] = \{m\}$ per ogni $m$ nel dominio di $\mathscr{M}$, pertanto questo caso si riduce essenzialmente al caso senza identità, come ci si può aspettare.

Teorema di Compattezza esso è anche soddisfacibile. Consideriamo pertanto un qualsiasi modello $\hat{\mathscr{M}} \models \hat{\Gamma}$. Abbiamo in particolare che $\hat{\mathscr{M}} \models \{d_i \neq d_j \mid i, j \in \mathbb{N}, i \neq j\}$. Pertanto $d_i^{\hat{\mathscr{M}}} \neq d_j^{\hat{\mathscr{M}}}$ per $i \neq j$, ovvero $\hat{\mathscr{M}}$ è definito su un insieme $\hat{M}$ contenente un numero infinito di elementi. Per ottenere un modello $\mathscr{M}$ per $\Gamma$ nel vecchio linguaggio $\mathscr{L}$ non resta altro che rimuovere da $\hat{\mathscr{M}}$ le funzioni di interpretazione delle nuove costanti (mantenendo però inalterato l'insieme $\hat{M}$ come dominio per $\mathscr{M}$). $\square$

**Corollario 4.5.** *La teoria dei gruppi finiti, la teoria degli anelli finiti, e la teoria dei campi finiti non sono assiomatizzabili al primordine.*

**Teorema 4.6** (Teorema di Skolem-Löwenheim verso il basso (caso numerabile))**.** *Sia* $\Gamma$ *un insieme soddisfacibile di* $\mathscr{L}$*-formule chiuse della logica con identità per* $\mathscr{L}$ *numerabile. Allora* $\Gamma$ *possiede un modello numerabile (finito o infinito).*

*Dimostrazione.* Sia $\Gamma$ soddisfacibile. Allora esso è anche finitamente soddisfacibile. Sulla base della dimostrazione del Teorema di Compattezza sappiamo che essendo $\mathscr{L}$ numerabile esiste un modello (infinito) *numerabile* $\mathscr{M}$ di $\Gamma_{\simeq} \cup AxCongr$ (ottenuto mediante il metodo della costruzione degli insiemi di Henkin). Per il Lemma 4.3 sappiamo che $[\mathscr{M}] \models \Gamma$. Poiché il numero delle classi di equivalenza sugli elementi di un insieme è ovviamente sempre minore o uguale del numero degli elementi dell'insieme, il dominio di $[\mathscr{M}]$ è finito o infinito numerabile. $\square$

Una conseguenza del Teorema di Skolem-Löwenheim verso il basso è che al primordine non è possibile in un linguaggio numerabile $\mathscr{L}$ trovare un sistema di assiomi che caratterizzi univocamente la struttura dei numeri reali. Infatti comunque identifichiamo un insieme $\Gamma$ di $\mathscr{L}$-formule tale che $\mathbb{R} \models \Gamma$, esiste sempre una struttura numerabile $\mathbb{R}'$ tale che $\mathbb{R}' \models \Gamma$.

Capitolo 5

# Teorie assiomatizzabili universalmente

## 5.1 Omomorfismi ed immersioni

**Definizione 5.1** (omomorfismo)**.** *Date $\mathscr{A}, \mathscr{B}$ $\mathscr{L}$-strutture con dominio $A$ e $B$ rispettivamente, una funzione $h : A \to B$ è un* omomorfismo *da $\mathscr{A}$ a $\mathscr{B}$ se:*

*(i)* $h(c^{\mathscr{A}}) = c^{\mathscr{B}}$ *per ogni $\mathscr{L}$-simbolo di costante $c$*

*(ii)* $h(f^{\mathscr{A}}(a_1, \ldots, a_n)) = f^{\mathscr{B}}(h(a_1), \ldots, h(a_n))$ *per ogni $\mathscr{L}$-simbolo di funzione $n$-ario e per ogni $a_1, \ldots, a_n \in A$*

*(iii)* $(a_1, \ldots, a_n) \in P^{\mathscr{A}} \Rightarrow (h(a_1), \ldots, h(a_n)) \in P^{\mathscr{B}}$ *per ogni $\mathscr{L}$-simbolo di predicato $n$-ario e per ogni $a_1, \ldots, a_n \in A$.*[1]

**Definizione 5.2** (immersione)**.** *Date $\mathscr{A}, \mathscr{B}$ $\mathscr{L}$-strutture con dominio $A$ e $B$ rispettivamente, una funzione $h : A \to B$ è un'*immersione *da $\mathscr{A}$ in $\mathscr{B}$ se è iniettiva, ed in aggiunta alle condizioni (i)-(iii) della Definizione 5.1 soddisfa anche la condizione inversa alla (iii), cioè:*

[1] La funzione $h : m \mapsto [m]$ che abbiamo considerato nella dimostrazione del Teorema di Compattezza per l'identità è quindi un esempio di omomorfismo.

*(iii')* $(h(a_1), ..., h(a_n)) \in P^{\mathscr{B}} \Rightarrow (a_1, ..., a_n) \in P^{\mathscr{A}}$ *per ogni simbolo di predicato $n$-ario e per ogni $a_1, ..., a_n \in A$.*[2]

**Definizione 5.3** (isomorfismo)**.** *Un* isomorfismo *è un'immersione biettiva.*

*Per due $\mathscr{L}$-strutture $\mathscr{A}, \mathscr{B}$ scriviamo $\mathscr{A} \approx \mathscr{B}$ quando esiste un isomorfismo tra di esse e diciamo che $\mathscr{A}$ e $\mathscr{B}$ sono* isomorfe.

Qualora l'immersione $h$ sia la funzione identità $id_{A,B} : A \to B, a \mapsto a$, per $A \subseteq B$, abbiamo che $\mathscr{A}$ è sottostruttura di $\mathscr{B}$:

**Definizione 5.4** (sottostruttura)**.** *Date $\mathscr{A}, \mathscr{B}$ $\mathscr{L}$-strutture con dominio $A$ e $B$ rispettivamente tali che $A \subseteq B$, $\mathscr{A}$ è una* sottostruttura *di $\mathscr{B}$ ($\mathscr{A} \subseteq \mathscr{B}$) se:*

*(i)* $c^{\mathscr{A}} = c^{\mathscr{B}}$ *per ogni $\mathscr{L}$-simbolo di costante $c$*

*(ii)* $f^{\mathscr{A}}(a_1, ..., a_n) = f^{\mathscr{B}}(a_1, ..., a_n)$ *per ogni simbolo di funzione $n$-ario e per ogni $a_1, ..., a_n \in A$*

*(iii)* $(a_1, ..., a_n) \in P^{\mathscr{A}} \Leftrightarrow (a_1, ..., a_n) \in P^{\mathscr{B}}$ *per ogni simbolo di predicato $n$-ario e per ogni $a_1, ..., a_n \in A$.*

**Lemma 5.5.** *Siano $\mathscr{A}, \mathscr{B}$ $\mathscr{L}$-strutture con domini $A$ e $B$ rispettivamente tali che $\mathscr{A} \subseteq \mathscr{B}$, e sia $\varphi(x_1, ..., x_n)$ una $\mathscr{L}$-formula senza quantificatori. Allora per ogni $a_1, .., a_n \in A$ si dà*

$$\mathscr{A} \models \varphi[a_1, ..., a_n] \Leftrightarrow \mathscr{B} \models \varphi[a_1, ..., a_n].$$

*Dimostrazione.* Notiamo innanzitutto per induzione sulla costruzione dei termini (esercizio) che per ogni $\mathscr{L}$-termine $t(x_1, ..., x_n)$ si ha

$$t^{\mathscr{A}}(a_1, ..., a_n) = t^{\mathscr{B}}(a_1, ..., a_n) \qquad (\S)$$

per ogni $a_1, .., a_n \in A$. La dimostrazione procede poi per induzione sulla costruzione delle formule (senza quantificatori). Vediamo solo il caso delle formule atomiche (i casi dei connettivi seguono banalmente per IH e sono lasciati per esercizio). Sia quindi $\varphi(x_1, ..., x_n)$ del tipo $P(t_1(x_1, ..., x_n), ..., t_m(x_1, ..., x_n))$ per $P$ $\mathscr{L}$-simbolo di predicato $m$-ario. Allora $\mathscr{A} \models P(t_1(x_1, ..., x_n), ..., t_m(x_1, ..., x_n))[a_1, .., a_n]$

[2]Se includessimo il simbolo di uguaglianza tra i simboli di predicato, allora l'iniettività di $h$ seguirebbe dalla condizione (iii').

sse $(t_1^{\mathscr{A}}(a_1,...,a_n),...,t_m^{\mathscr{A}}(a_1,...,a_n)) \in P^{\mathscr{A}}$
sse $(t_1^{\mathscr{A}}(a_1,...,a_n),...,t_m^{\mathscr{A}}(a_1,...,a_n)) \in P^{\mathscr{B}}$ (per definizione di sottostruttura) sse $(t_1^{\mathscr{B}}(a_1,...,a_n),...,t_m^{\mathscr{B}}(a_1,...,a_n)) \in P^{\mathscr{B}}$ (per §)
sse $\mathscr{B} \models P(t_1(x_1,...,x_n),...,t_m(x_1,...,x_n))[a_1,..,a_n]$.

Il caso di $\varphi(x_1,...,x_n) \equiv t_1(x_1,...,x_n) = t_2(x_1,...,x_n)$ è analogo. □

**Corollario 5.6** (del Sottomodello)**.** *Siano $\mathscr{A}$, $\mathscr{B}$ $\mathscr{L}$-strutture (con domini A e B rispettivamente) tali che $\mathscr{A} \subseteq \mathscr{B}$, e sia $\varphi$ una $\mathscr{L}$-formula chiusa universale, ovvero del tipo $(\forall x_1)...(\forall x_n)\psi(x_1,...,x_n)$, dove la formula $\psi(x_1,...,x_n)$ è senza quantificatori. Se $\mathscr{B} \models \varphi$ allora $\mathscr{A} \models \varphi$.*

*Dimostrazione.* Supponiamo

$$\mathscr{B} \models (\forall x_1)...(\forall x_n)\psi(x_1,...,x_n).$$

Allora $\mathscr{B} \models \psi[b_1,...,b_n]$ per tutti i $b_1,...,b_n \in B$. In particolare $\mathscr{B} \models \psi[a_1,...,a_n]$ per tutti gli $a_1,...,a_n \in A$. Per il Lemma 5.5 abbiamo che $\mathscr{A} \models \psi[a_1,...,a_n]$ per tutti gli $a_1,...,a_n \in A$, da cui

$$\mathscr{A} \models (\forall x_1)...(\forall x_n)\psi(x_1,...,x_n).$$

□

**Notazione 5.7.** *Data una teoria $T$ indichiamo con $T_\forall$ l'insieme delle formule (chiuse) universali di $T$.*

**Proposizione 5.8.** *Sia data una teoria $T$. Se $T$ è assiomatizzabile universalmente (ovvero $T_\forall \vdash T$), allora $T$ si conserva per sottostrutture, ovvero $\mathscr{M}_1 \models T$ & $\mathscr{M}_2 \subseteq \mathscr{M}_1 \Rightarrow \mathscr{M}_2 \models T$.*

*Dimostrazione.* È una conseguenza del Teorema di Validità. Sia infatti $\mathscr{M}_1 \models T$ e $\mathscr{M}_2 \subseteq \mathscr{M}_1$. Giacché $T_\forall \subseteq T$, allora $\mathscr{M}_1 \models T_\forall$. Per il Corollario 5.6 segue che $\mathscr{M}_2 \models T_\forall$. Poiché per ipotesi $T_\forall \vdash T$, allora $T_\forall \models T$ (Teorema di Validità), da cui $\mathscr{M}_2 \models T$. □

**Lemma 5.9.** *Siano $\mathscr{A}$, $\mathscr{B}$ $\mathscr{L}$-strutture con domini A e B rispettivamente e sia $h : A \to B$ un isomorfismo tra $\mathscr{A}$ e $\mathscr{B}$. Sia $\varphi(x_1,...,x_n)$ una $\mathscr{L}$-formula. Allora per ogni $a_1,..,a_n \in A$ si dà*

$$\mathscr{A} \models \varphi[a_1,...,a_n] \Leftrightarrow \mathscr{B} \models \varphi[h(a_1),...,h(a_n)].$$

*Dimostrazione.* Notiamo innanzitutto per induzione sulla costruzione dei termini (esercizio) che, per il solo fatto che $h$ è un omomorfismo, per ogni $\mathscr{L}$-termine $t(x_1,...,x_n)$ si ha

$$h(t^{\mathscr{A}}(a_1,...,a_n)) = t^{\mathscr{B}}(h(a_1),...,h(a_n)) \qquad (\S\S).$$

La dimostrazione procede poi per induzione sulla costruzione delle formule. Vediamo il caso delle formule atomiche e quello dei quantificatori, mentre quello dei connettivi è lasciato per esercizio.

Sia $\varphi(x_1,...,x_n) \equiv P(t_1(x_1,...,x_n),...,t_m(x_1,...,x_n))$ per $P$ $\mathscr{L}$-simbolo di predicato.

Allora $\mathscr{A} \models P(t_1(x_1,...,x_n),...,t_m(x_1,...,x_n))[a_1,..,a_n]$
sse $(t_1^{\mathscr{A}}(a_1,...,a_n),...,t_m^{\mathscr{A}}(a_1,...,a_n)) \in P^{\mathscr{A}}$
sse $(h(t_1^{\mathscr{A}}(a_1,...,a_n)),...,h(t_m^{\mathscr{A}}(a_1,...,a_n))) \in P^{\mathscr{B}}$ (perché $h$ è un'immersione) sse $(t_1^{\mathscr{B}}(h(a_1),...,h(a_n)),...,t_m^{\mathscr{B}}(h(a_1),...,h(a_n))) \in P^{\mathscr{B}}$ (per §§)
sse $\mathscr{B} \models P(t_1(x_1,...,x_n),...,t_m(x_1,...,x_n))[h(a_1),..,h(a_n)]$.

Sia $\varphi(x_1,...,x_n) \equiv t_1(x_1,...,x_n) = t_2(x_1,...,x_n)$.

Allora $\mathscr{A} \models t_1(x_1,...,x_n) = t_2(x_1,...,x_n)[a_1,..,a_n]$ sse $t_1^{\mathscr{A}}(a_1,...,a_n) = t_2^{\mathscr{A}}(a_1,...,a_n)$ sse $h(t_1^{\mathscr{A}}(a_1,...,a_n)) = h(t_2^{\mathscr{A}}(a_1,...,a_n))$ (la direzione *solo se* si dà perché $h$ è una funzione, quella *se* perché $h$ è iniettiva) sse $t_1^{\mathscr{B}}(h(a_1),...,h(a_n)) = t_2^{\mathscr{B}}(h(a_1),...,h(a_n))$ (per §§)
sse $\mathscr{B} \models t_1(x_1,...,x_n) = t_2(x_1,...,x_n)[h(a_1),..,h(a_n)]$.

Sia $\varphi(x_1,...,x_n) \equiv (\forall x)\psi(x,x_1,...,x_n)$.

Allora $\mathscr{A} \models (\forall x)\psi(x,x_1,...,x_n)[a_1,...,a_n]$ sse $\mathscr{A} \models \psi[a,a_1,...,a_n]$ per ogni $a \in A$ sse $\mathscr{B} \models \psi[h(a),h(a_1),...,h(a_n)]$ per ogni $a \in A$ (IH) sse $\mathscr{B} \models \psi[b,h(a_1),...,h(a_n)]$ per ogni $b \in B$ (la direzione *solo se* deriva dal fatto che $h$ è suriettiva, quella *se* è ovvia)
sse $\mathscr{B} \models (\forall x)\psi(x,x_1,...,x_n)[h(a_1),...,h(a_n)]$.

Sia $\varphi(x_1,...,x_n) \equiv (\exists x)\psi(x,x_1,...,x_n)$.

Allora $\mathscr{A} \models (\exists x)\psi(x,x_1,...,x_n)[a_1,...,a_n]$ sse $\mathscr{A} \models \psi[a,a_1,...,a_n]$ per un $a \in A$ sse $\mathscr{B} \models \psi[h(a),h(a_1),...,h(a_n)]$ per un $a \in A$ (IH) sse $\mathscr{B} \models \psi[b,h(a_1),...,h(a_n)]$ per un $b \in B$ (la direzione *solo se* è ovvia, quella *se* deriva dal fatto che $h$ è suriettiva)
sse $\mathscr{B} \models (\exists x)\psi(x,x_1,...,x_n)[h(a_1),...,h(a_n)]$. □

**Corollario 5.10.** *Siano $\mathscr{A},\mathscr{B}$ $\mathscr{L}$-strutture isomorfe. Allora per ogni $\mathscr{L}$-formula chiusa $\varphi$ si dà $\mathscr{A} \models \varphi \Leftrightarrow \mathscr{B} \models \varphi$.*

**Definizione 5.11** (Equivalenza elementare)**.** *Date due $\mathscr{L}$-strutture $\mathscr{A}$ e $\mathscr{B}$, diciamo che $\mathscr{A}$ è* elementarmente equivalente *a $\mathscr{B}$ ($\mathscr{A} \equiv \mathscr{B}$) se per ogni $\mathscr{L}$-formula chiusa $\varphi$ si dà $\mathscr{A} \models \varphi \Leftrightarrow \mathscr{B} \models \varphi$.*

**Osservazione 5.12.** *Segue ovviamente dal Corollario 5.10 che* $\mathscr{A} \approx \mathscr{B} \Rightarrow \mathscr{A} \equiv \mathscr{B}$. *L'inverso tuttavia non vale necessariamente. Consideriamo infatti la struttura* $\mathbb{R}$ *dei numeri reali in un linguaggio* $\mathscr{L}$*-numerabile e sia* $Ver(\mathbb{R}) := \{\varphi \mid \varphi$ *è* $\mathscr{L}$*-formula chiusa e* $\mathbb{R} \models \varphi\}$. *Allora per il Teorema 4.6 di Skolem-Löwenheim verso il basso esiste una* $\mathscr{L}$*-struttura numerabile* $\mathbb{R}'$ *tale che* $\mathbb{R}' \models Ver(\mathbb{R})$. *Inoltre, per ogni formula chiusa* $\varphi$ *si ha che* $\mathbb{R} \models \varphi \Longleftrightarrow \mathbb{R}' \models \varphi$ *(esercizio; conseguenza del fatto che in qualsiasi* $\mathscr{L}$*-struttura* $\mathscr{A}$ *si ha* $\mathscr{A} \models \varphi$ *oppure* $\mathscr{A} \models \neg\varphi$*). Perciò* $\mathbb{R}$ *è elementarmente equivalente a* $\mathbb{R}'$. *Tuttavia ovviamente* $\mathbb{R}$ *e* $\mathbb{R}'$ *non possono essere isomorfe perché le loro cardinalità sono differenti.*

## *5.2 Diagrammi*

**Notazione 5.13** (Letterali)**.** *Un* letterale *in un linguaggio* $\mathscr{L}$ *è una* $\mathscr{L}$*-formula atomica o la negazione di una* $\mathscr{L}$*-formula atomica.*

**Definizione 5.14** (Diagrammi)**.** *Sia data una* $\mathscr{L}$*-struttura* $\mathscr{A}$ *definita sull'insieme A. Consideriamo il linguaggio esteso* $\mathscr{L}_A$ *che aggiunge ad* $\mathscr{L}$ *una nuova costante* $\overline{a}$ *per ogni* $a \in A$. *Definiamo allora il* diagramma *di* $\mathscr{A}$ *come l'insieme*

$$\mathrm{Diag}(\mathscr{A}) := \{L(\overline{a}_1/x_1,\ldots,\overline{a}_n/x_n) \mid L(x_1,\ldots,x_n) \text{ è un } \mathscr{L}\text{-letterale}$$

$$\text{e } \mathscr{A} \models L(x_1,\ldots,x_n)[a_1,..,a_n]\}$$

Alcune osservazioni sulla notazione. Con "$L(x_1,\ldots,x_n)$" intendiamo dire che le variabili libere del letterale $L$ sono in $\{x_1,\ldots,x_n\}$. Siccome nel letterale $L(\overline{a}_1/x_1,\ldots,\overline{a}_n/x_n)$ queste variabili, se originariamente presenti, sono state rimpiazzate mediante nuove costanti di $\mathscr{L}_A$, il letterale così ottenuto è ovviamente una formula chiusa, ma non è detto che $L(\overline{a}_1/x_1,\ldots,\overline{a}_n/x_n)$ non possa contenere anche costanti del vecchio linguaggio $\mathscr{L}$. Il fatto è che in quello che segue noi siamo interessati alle sostituzioni che riguardano solo le nuove costanti, e quindi solo su queste ci concentriamo. Ad esempio, la formula atomica $P(\overline{a},c)$, per $c$ costante di $\mathscr{L}$, può essere vista come ottenuta da $P(x,y)$ rimpiazzando $x$ con $\overline{a}$ ed $y$ con $c$ oppure come ottenuta da $P(x,c)$ rimpiazzando $x$ con $\overline{a}$. Nei seguenti risultati noi ci atterremo a questa seconda lettura.

**Esempio 5.15.** *Per* $\langle \mathbb{N}, 0, 1, +, \cdot \rangle$ *e* $L(x, y) \equiv 1 + x = y$ *abbiamo* $1 + \overline{5} = \overline{6} \in \mathrm{Diag}(\mathbb{N})$, $1 + \overline{5} = \overline{10} \notin \mathrm{Diag}(\mathbb{N})$. *Per* $L(x, y, z) \equiv \neg\, x \cdot y = z$ *abbiamo* $\neg\, \overline{5} \cdot \overline{2} = \overline{7} \in \mathrm{Diag}(\mathbb{N})$, $\neg\, \overline{5} \cdot \overline{2} = \overline{10} \notin \mathrm{Diag}(\mathbb{N})$.

Si osservi che anche qualora $\mathscr{L}$ sia numerabile, $\mathscr{L}_A$ può ben non esserlo!

**Definizione 5.16.** *Siano* $\mathscr{A}$ *e* $\mathscr{B}$ *due* $\mathscr{L}$*-strutture con dominio A e B rispettivamente e sia* $h : A \to B$ *una qualsiasi funzione. Indichiamo con* $\mathscr{B}^h_{\mathscr{A}}$ *la* $\mathscr{L}_A$*-struttura che arricchisce* $\mathscr{B}$ *con le funzioni di interpretazione delle nuove costanti di* $\mathscr{L}_A$ *non in* $\mathscr{L}$ *in modo che* $\overline{a}^{\mathscr{B}^h_{\mathscr{A}}} := h(a)$.

**Lemma 5.17** (del Diagramma)**.** *Siano* $\mathscr{A}$ *e* $\mathscr{B}$ *due* $\mathscr{L}$*-strutture con dominio A e B rispettivamente e sia* $h : A \to B$ *una qualsiasi funzione. Se* $\mathscr{B}^h_{\mathscr{A}} \models \mathrm{Diag}(\mathscr{A})$, *allora h è un isomorfismo tra* $\mathscr{A}$ *ed una sottostruttura* $\mathscr{A}^*$ *di* $\mathscr{B}$.

*Dimostrazione.* È sufficiente dimostrare che se vale $\mathscr{B}^h_{\mathscr{A}} \models \mathrm{Diag}(\mathscr{A})$, allora $h$ è un'immersione di $\mathscr{A}$ in $\mathscr{B}$. Avremo che la desiderata $\mathscr{A}^*$ è la sottostruttura di $\mathscr{B}$ avente come dominio $h(A) \subseteq B$, cioè l'immagine di $A$ in $B$ rispetto ad $h$.

Sia quindi $\mathscr{B}^h_{\mathscr{A}} \models \mathrm{Diag}(\mathscr{A})$. Allora:

(i) $h(c^{\mathscr{A}}) = c^{\mathscr{B}}$ per ogni $\mathscr{L}$-costante $c$: sia infatti $c^{\mathscr{A}} := a \in A$. Allora $c = \overline{a} \in \mathrm{Diag}(\mathscr{A})$ e quindi per ipotesi $\mathscr{B}^h_{\mathscr{A}} \models c = \overline{a}$, da cui $c^{\mathscr{B}^h_{\mathscr{A}}} = \overline{a}^{\mathscr{B}^h_{\mathscr{A}}}$. Ricordando che $\overline{a}^{\mathscr{B}^h_{\mathscr{A}}} := h(a)$, abbiamo che $c^{\mathscr{B}^h_{\mathscr{A}}} = h(a)$. Ma $c^{\mathscr{B}^h_{\mathscr{A}}} = c^{\mathscr{B}}$, dunque $h(c^{\mathscr{A}}) = h(a) = c^{\mathscr{B}}$;

(ii) $h(f^{\mathscr{A}}(a_1, .., a_n)) = f^{\mathscr{B}}(h(a_1), ..., h(a_n))$ per ogni $\mathscr{L}$-simbolo di funzione $n$-ario $f$ e $a_1, ..., a_n \in A$: sia infatti $f^{\mathscr{A}}(a_1, ..., a_n) := a \in A$. Allora $f(\overline{a}_1, ..., \overline{a}_n) = \overline{a} \in \mathrm{Diag}(\mathscr{A})$ e quindi per ipotesi abbiamo che $\mathscr{B}^h_{\mathscr{A}} \models f(\overline{a}_1, ..., \overline{a}_n) = \overline{a}$, da cui $f^{\mathscr{B}^h_{\mathscr{A}}}(\overline{a}_1^{\mathscr{B}^h_{\mathscr{A}}}, ..., \overline{a}_n^{\mathscr{B}^h_{\mathscr{A}}}) = \overline{a}^{\mathscr{B}^h_{\mathscr{A}}}$. Ricordando che $\overline{a}_i^{\mathscr{B}^h_{\mathscr{A}}} := h(a_i)$, per $1 \le i \le n$, $\overline{a}^{\mathscr{B}^h_{\mathscr{A}}} := h(a)$ e $f^{\mathscr{B}^h_{\mathscr{A}}} = f^{\mathscr{B}}$, abbiamo che $f^{\mathscr{B}}(h(a_1), ..., h(a_n)) = h(a) = h(f^{\mathscr{A}}(a_1, ..., a_n))$;

(iii)+(iii') $(a_1, ..., a_n) \in P^{\mathscr{A}} \Leftrightarrow (h(a_1), ..., h(a_n)) \in P^{\mathscr{B}}$ per ogni $\mathscr{L}$-simbolo di predicato $n$-ario $P$ e $a_1, ..., a_n \in A$:

$\Rightarrow$) sia infatti $(a_1, .., a_n) \in P^{\mathscr{A}}$. Allora $P(\overline{a}_1, ..., \overline{a}_n) \in \mathrm{Diag}(\mathscr{A})$ e quindi per ipotesi $\mathscr{B}^h_{\mathscr{A}} \models P(\overline{a}_1, ..., \overline{a}_n)$, da cui $(\overline{a}_1^{\mathscr{B}^h_{\mathscr{A}}}, ..., \overline{a}_n^{\mathscr{B}^h_{\mathscr{A}}}) \in P^{\mathscr{B}^h_{\mathscr{A}}}$. Ricordando che $\overline{a}_i^{\mathscr{B}^h_{\mathscr{A}}} := h(a_i)$ per $1 \leq i \leq n$ e $P^{\mathscr{B}^h_{\mathscr{A}}} = P^{\mathscr{B}}$, abbiamo che $(h(a_1), ..., h(a_n)) \in P^{\mathscr{B}}$;

$\Leftarrow$) supponiamo che $(a_1, ..., a_n) \notin P^{\mathscr{A}}$. Allora per definizione di diagramma $\neg P(\overline{a}_1, ..., \overline{a}_n) \in \mathrm{Diag}(\mathscr{A})$ e quindi per ipotesi abbiamo che $\mathscr{B}^h_{\mathscr{A}} \models \neg P(\overline{a}_1, ..., \overline{a}_n)$, cioè $\mathscr{B}^h_{\mathscr{A}} \not\models P(\overline{a}_1, ..., \overline{a}_n)$, da cui segue che $(\overline{a}_1^{\mathscr{B}^h_{\mathscr{A}}}, ..., \overline{a}_n^{\mathscr{B}^h_{\mathscr{A}}}) \notin P^{\mathscr{B}^h_{\mathscr{A}}}$. Ricordando che $\overline{a}_i^{\mathscr{B}^h_{\mathscr{A}}} := h(a_i)$ per $1 \leq i \leq n$ e $P^{\mathscr{B}^h_{\mathscr{A}}} = P^{\mathscr{B}}$, abbiamo $(h(a_1), ..., h(a_n)) \notin P^{\mathscr{B}}$.

Adattando l'argomento per (iii') al caso del simbolo di uguaglianza = (esercizio) possiamo dedurre che $h$ è iniettiva. Pertanto $h$ è effettivamente un'immersione. □

**Lemma 5.18.** *Sia $T$ una $\mathscr{L}$-teoria soddisfacibile e sia $\mathscr{M} \models T_\forall$ per una qualche $\mathscr{L}$-struttura $\mathscr{M}$. Allora $T \cup \mathrm{Diag}(\mathscr{M})$ è soddisfacibile.*

*Dimostrazione.* Supponiamo che $T \cup \mathrm{Diag}(\mathscr{M})$ non sia soddisfacibile (come insieme di $\mathscr{L}_M$-formule, per $M$ dominio di $\mathscr{M}$). Allora per Teorema di Compattezza (per ogni cardinalità) esistono $\mathscr{L}_M$-letterali chiusi $L_1(\overline{m}_1/x_1, ..., \overline{m}_n/x_n), ..., L_k(\overline{m}_1/x_1, ..., \overline{m}_n/x_n) \in \mathrm{Diag}(\mathscr{M})$, per $m_1, ... m_n \in M$, tali che

$$T \cup \{L_1(\overline{m}_1/x_1, ..., \overline{m}_n/x_n), ..., L_k(\overline{m}_1/x_1, ..., \overline{m}_n/x_n)\}$$

non è soddisfacibile. Questo significa che anche

$$T \cup \{\wedge_{1 \leq i \leq k} L_i(\overline{m}_1/x_1, ..., \overline{m}_n/x_n)\}$$

è insoddisfacibile, e quindi $T \models \neg \wedge_{1 \leq i \leq k} L_i(\overline{m}_1/x_1, ..., \overline{m}_n/x_n)$ (esercizio). Mostriamo che allora vale anche $T \models (\forall x_1)...(\forall x_n) \neg \wedge_{1 \leq i \leq k} L_i(x_1, ..., x_n)$. Sia infatti $\mathscr{M}' \models T$ per una qualsiasi $\mathscr{L}$-struttura $\mathscr{M}'$ con dominio $M'$. Siano $m'_1, ... m'_n$ elementi arbitrari di $M'$. Poiché in $T$ non compaiono le nuove costanti di $\mathscr{L}_M$, se consideriamo la $\mathscr{L}_M$-struttura $\mathscr{M}''$ che è in tutto uguale ad $\mathscr{M}'$ fatto salvo che $\overline{m}_i^{\mathscr{M}''} := m'_i$, per $1 \leq i \leq n$, questa è ancora un modello di $T$ (visto nel linguaggio $\mathscr{L}_M$). Per definizione di conseguenza logica segue che $\mathscr{M}''$ è anche un modello di $\neg \wedge_{1 \leq i \leq k} L_i(\overline{m}_1/x_1, ..., \overline{m}_n/x_n)$, cioè $\mathscr{M}'' \models \neg \wedge_{1 \leq i \leq k} L_i(\overline{m}_1/x_1, ..., \overline{m}_n/x_n)$, e quindi $\mathscr{M}'' \models \neg \wedge_{1 \leq i \leq k} L_i(x_1, ..., x_n)[m'_1, ..., m'_n]$, ma allora anche $\mathscr{M}' \models \neg \wedge_{1 \leq i \leq k} L_i(x_1, ..., x_n)[m'_1, ..., m'_n]$. Concludiamo $\mathscr{M}' \models (\forall x_1)...(\forall x_n) \neg \wedge_{1 \leq i \leq k} L_i(x_1, ..., x_n)$. Allora per definizione di conseguenza logica $T \models (\forall x_1)...(\forall x_n) \neg \wedge_{1 \leq i \leq k} L_i(x_1, ..., x_n)$, e, per il Teorema di Completezza, $T \vdash (\forall x_1)...(\forall x_n)\neg \wedge_{1\leq i\leq k} L_i(x_1,...,x_n)$. Questo indica che $(\forall x_1)...(\forall x_n)\neg \wedge_{1\leq i\leq k} L_i(x_1,...,x_n) \in T_\forall$. Giacché $\mathscr{M} \models T_\forall$, allora $\mathscr{M} \models (\forall x_1)...(\forall x_n)\neg \wedge_{1\leq i\leq k} L_i(x_1,...,x_n)$, e pertanto $\mathscr{M} \not\models \wedge_{1\leq i\leq k} L_i(x_1,...,x_n)[m_1,..,m_n]$. Al contempo abbiamo anche che $\mathscr{M} \models \wedge_{1\leq i\leq k} L_i(x_1,...,x_n)[m_1,...,m_n]$ in quanto per ipotesi $L_i(\overline{m}_1/x_1,...,\overline{m}_n/x_n) \in \mathrm{Diag}(\mathscr{M})$ per $1 \leq i \leq k$. Contraddizione. □

**Teorema 5.19** (delle Teorie Universali)**.** *Sia $T$ una $\mathscr{L}$-teoria tale che $\mathscr{M}_1 \models T$ & $\mathscr{M}_2 \subseteq \mathscr{M}_1 \Rightarrow \mathscr{M}_2 \models T$. Allora $T$ è assiomatizzabile con formule universali.*

*Dimostrazione.* Dimostreremo che $T_\forall \vdash T$.

Se $T$ è insoddisfacibile, la cosa è banale (basta osservare che per Teorema di Compattezza, caso generale, $(\forall x)\bot \in T_\forall$). Sia ora invece $T$ soddisfacibile. Per il Teorema di Completezza è sufficiente dimostrare che $T_\forall \models T$. Sia quindi $\mathscr{M} \models T_\forall$ per una $\mathscr{L}$-struttura $\mathscr{M}$ con dominio $M$. Per il Lemma 5.18, $T \cup \mathrm{Diag}(\mathscr{M})$ è soddisfacibile. Sia quindi $\mathscr{M}_1'$ un $\mathscr{L}_M$-modello di $T \cup \mathrm{Diag}(\mathscr{M})$. Consideriamo ora la $\mathscr{L}$-struttura $\mathscr{M}_1$ che si ottiene ripulendo $\mathscr{M}_1'$ da tutte le funzioni di interpretazione delle nuove costanti di $\mathscr{L}_M$. Siccome in $T$ non compaiono le nuove costanti di $\mathscr{L}_M$, abbiamo che $\mathscr{M}_1$ è un $\mathscr{L}$-modello di $T$. Ovviamente i domini di $\mathscr{M}_1'$ e $\mathscr{M}_1$ coincidono. Sia $M_1$ tale dominio. Prendiamo ora la funzione $h : M \to M_1$ tale che $h(m) := \overline{m}^{\mathscr{M}_1'}$. Allora $\mathscr{M}_1' = (\mathscr{M}_1)^h_{\mathscr{M}}$. Per il Lemma 5.17 del Diagramma, $h$ è un isomorfismo tra $\mathscr{M}$ ed una sottostruttura $\mathscr{M}_2$ di $\mathscr{M}_1$ (cioè $\mathscr{M}_2 = \mathscr{M}^*$). Per ipotesi $\mathscr{M}_2 \models T$. Ma come abbiamo detto $\mathscr{M} \approx \mathscr{M}_2$, e quindi $\mathscr{M} \equiv \mathscr{M}_2$ ($\mathscr{M}$ ed $\mathscr{M}_2$ sono elementarmente equivalenti) per l'Osservazione 5.12. Ne segue che $\mathscr{M} \models T$. Concludiamo che $T_\forall \models T$. □

**Corollario 5.20.** *Una teoria $T$ è assiomatizzabile universalmente sse si conserva per sottostrutture.*

*Dimostrazione.* Segue dalla Proposizione 5.8 e dal Teorema 5.19. □

**Teorema 5.21.** *Una $\mathscr{L}$-formula chiusa $\varphi$ soddisfacibile è equivalente ad una $\mathscr{L}$-formula chiusa universale sse si conserva per sottostrutture, ovvero qualora $\mathscr{A} \subseteq \mathscr{B}$ & $\mathscr{B} \models \varphi \Rightarrow \mathscr{A} \models \varphi$.*

*Dimostrazione.* $\Rightarrow$) Segue dal Corollario 5.6 del Sottomodello.

$\Leftarrow$) Innanzitutto, con dimostrazione simile a quella della Proposizione 5.8, si può vedere che la teoria $T$ assiomatizzata da $\varphi$ si conserva per sottostrutture esattamente come $\varphi$ (esercizio). Per il Teorema 5.19 delle Teorie Universali avremo quindi che $T_\forall \vdash T$. Giacché $\varphi \in T$, si ha pertanto che $T_\forall \vdash \varphi$, esistono cioè $\psi_1, \ldots, \psi_n \in T_\forall$ tali che $\{\psi_1, \ldots, \psi_n\} \vdash \varphi$. Si osservi che $\psi_1 \wedge \ldots \wedge \psi_n$ è equivalente ad una formula universale $\psi$ (basta considerare una forma normale prenessa di $\psi_1 \wedge \ldots \wedge \psi_n$, vedasi sotto). Da un lato abbiamo $\varphi \vdash \psi$ in quanto $\psi \in T_\forall$, $T_\forall \subseteq T$ e $\varphi \vdash T$, dall'altro $\psi \vdash \varphi$ in quanto $\{\psi_1, \ldots, \psi_n\} \vdash \varphi$. Per teorema di Validità $\varphi$ è logicamente equivalente a $\psi$. $\square$

Esempi di teorie non assiomatizzabili universalmente sono la teoria dei monoidi nel linguaggio $\{\dot{+}\}$, la teoria dei gruppi nel linguaggio $\{\dot{0}, \dot{+}\}$, e la teoria dei campi nel linguaggio $\{\dot{0}, \dot{1}, \dot{-}, \dot{+}, *\}$. Ad esempio, l'assioma (k2) affermante l'esistenza degli elementi inversi rispetto a $*$ non è equivalente ad una formula universale in quanto non si conserva per sottostrutture: $\mathbb{Z}$ è una sottostruttura di $\mathbb{Q}$ per la quale gli elementi inversi rispetto a $\times$ non esistono (eccetto che per 1).

Tuttavia sono assiomatizzabili universalmente la teoria dei monoidi nel linguaggio $\{\dot{0}, \dot{+}\}$, la teoria dei gruppi nel linguaggio $\{\dot{0}, \dot{-}, \dot{+}\}$, e la teoria dei campi nel linguaggio $\{\dot{0}, \dot{1}, \dot{-}, {}^{-1}, \dot{+}, *\}$. Per tornare al caso dei campi, si noti che in questo caso l'esempio di $\mathbb{Z}$ e $\mathbb{Q}$ in effetti viene a cadere, in quanto $\mathbb{Z}$ non può più essere sottostruttura di $\mathbb{Q}$ rispetto a tale linguaggio. Ed infatti l'assioma (k2') ha la forma di una formula universale.

Questo fatto ci suggerisce l'idea che tutte le teorie siano assiomatizzabili universalmente scegliendo un apposito linguaggio che consenta "universalizzare" gli assiomi contenenti quantificatori esistenziali. Ci occuperemo di questo nella prossima sezione.

## *5.3 Teoremi di Skolem ed Herbrand*

Sappiamo che ogni formula $\varphi$ (contenente quantificatori) è equivalente ad una formula in *forma normale prenessa*, ovvero una formula del tipo $(Qx_1)\ldots(Qx_n)\psi$ dove $Q \in \{\forall, \exists\}$, per $1 \leq i \leq n$, e $\psi$ è una formula senza quantificatori. Per rendere $\varphi$ in forma normale prenessa è per prima cosa opportuno ridenominare le variabili vincolate in modo da ottenere una *formula rettificata*. Questa è una formula nella quale ciascun quantificatore contiene variabili diverse da quelle degli altri ed inoltre non esistono variabili che abbiano sia occorrenze libere che occorrenti in un quantificatore (spesso noi operiamo con formule chiuse, nel qual caso questa seconda condizione è banalmente già soddisfatta dal principio). La ridenominazione delle variabili vincolate può essere fatta utilizzando le equivalenze del tipo

$$\vdash (Qx)\chi \leftrightarrow (Qy)\xi$$

con $Q \in \{\forall, \exists\}$. Sia quindi $\varphi'$ una forma rettificata di $\varphi$. Mettiamo ora $\varphi'$ in forma normale prenessa utilizzando le equivalenze (sempre dimostrabili per formule rettificate)

$$\vdash \neg(Qx)\chi \leftrightarrow (\overline{Q}x)\neg\chi,$$

$$\vdash (\chi\clubsuit(Qx)\xi) \leftrightarrow (Qx)(\chi\clubsuit\xi), \quad \vdash ((Qx)\chi\clubsuit\xi) \leftrightarrow (Qx)(\chi\clubsuit\xi)$$

$$\vdash (\chi \to (Qx)\xi) \leftrightarrow (Qx)(\chi \to \xi), \quad \vdash ((Qx)\chi \to \xi) \leftrightarrow (\overline{Q}x)(\chi \to \xi),$$

per $\clubsuit \in \{\wedge, \vee\}$, $Q \in \{\forall, \exists\}$, e $\overline{Q} \in \{\forall, \exists\} \setminus \{Q\}$.

Otteniamo così $\varphi''$. Dimostriamo ora il seguente risultato generale:

**Teorema 5.22** (Teorema di Skolem)**.** *Sia* $(\forall x_1)\dots(\forall x_n)(\exists y)\psi$ *una* $\mathscr{L}$*-formula chiusa in forma normale prenessa rettificata (con* $\psi$ *possibilmente contenente quantificatori). Sia* $f$ *un simbolo di funzione* $n$*-ario non contenuto in* $\mathscr{L}$*. Allora* $(\forall x_1)\dots(\forall x_n)\psi(x_1,\dots,x_n, f(x_1,\dots,x_n)/y)$ *è soddisfacibile sse* $(\forall x_1)\dots(\forall x_n)(\exists y)\psi$ *è soddisfacibile.*

*Dimostrazione.* $\Rightarrow$) Sia $\mathscr{M} \models (\forall x_1)\dots(\forall x_n)\psi(x_1,\dots,x_n, f(x_1,\dots,x_n)/y)$ per $\mathscr{M}$ $\mathscr{L}\cup\{f\}$-struttura con dominio $M$. Allora per ogni $m_1,\dots,m_n \in M$ si dà

$$\mathscr{M} \models \psi(x_1,\dots,x_n,y)[m_1,\dots,m_n, f^{\mathscr{M}}(m_1,\dots,m_n)],$$

per cui per ogni $m_1,\dots,m_n \in M$ esiste $m \in M$ tale che

$$\mathscr{M} \models \psi(x_1,\dots,x_n,y)[m_1,\dots,m_n,m]$$

(basta porre $m := f^{\mathscr{M}}(m_1,\dots,m_n) \in M$). Perciò

$$\mathscr{M} \models (\forall x_1)\dots(\forall x_n)(\exists y)\psi,$$

e quindi $\mathscr{M}$ è un $\mathscr{L} \cup \{f\}$-modello di $(\forall x_1)...(\forall x_n)(\exists y)\psi$. Per ottenere un semplice $\mathscr{L}$-modello $\mathscr{M}'$ di tale formula è sufficiente cancellare in $\mathscr{M}$ l'interpretazione di $f$.

$\Leftarrow$) Sia $\mathscr{M} \models (\forall x_1)...(\forall x_n)(\exists y)\psi$ per un qualche $\mathscr{L}$-modello $\mathscr{M}$ con dominio $M$. Questo significa che per ogni $m_1,...,m_n \in M$ esiste $m \in M$ tale che $\mathscr{M} \models \psi[m_1,...,m_n,m]$. Per ogni $m_1,...,m_n \in M$ consideriamo quindi l'insieme (non vuoto!)

$$I(m_1,...,m_n) := \left\{ m \in M \mid \mathscr{M} \models \psi[m_1,...,m_n,m] \right\}.$$

Tramite l'assioma di scelta selezioniamo un elemento $m_{m_1,...,m_n} \in I(m_1,...,m_n)$ e definiamo la funzione

$$F : M^n \to M, (m_1,...,m_n) \mapsto m_{m_1,...,m_n}.$$

Quindi si dà $\mathscr{M} \models \psi[m_1,...,m_n,F(m_1,...,m_n)]$ per ogni $m_1,...,m_n \in M$. Estendiamo l'$\mathscr{L}$-struttura $\mathscr{M}$ ad una $\mathscr{L} \cup \{f\}$-struttura $\mathscr{M}'$ ponendo $f^{\mathscr{M}'} := F$. Ovviamente si dà

$$\mathscr{M}' \models \psi(x_1,...,x_n, f(x_1,...,x_n)/y)[m_1,...,m_n]$$

per ogni $m_1,...,m_n \in M$, da cui

$$\mathscr{M}' \models (\forall x_1)...(\forall x_n)\psi(x_1,...,x_n, f(x_1,...,x_n)/y).$$

$\square$

**Osservazione 5.23.** *Trattiamo le formule del tipo* $(\exists y)\psi$ *come caso particolare, sostituendo* $y$ *in* $\psi$ *mediante una nuova funzione* $0$*-aria, cioè una costante.*

Per ottenere quindi una *forma di Skolem* della nostra formula normale prenessa rettificata $\varphi''$ menzionata sopra rimpiazziamo ogni quantificatore esistenziale in essa, partendo dal più esterno fino a raggiungere il più interno muovendosi verso destra, utilizzando ogni volta un nuovo simbolo di funzione non originariamente presente in $\varphi''$ e non utilizzato in qualche passo di rimpiazzamento precedente. Quella che otteniamo alla fine sarà una formula soddisfacibile in strutture funzionalemente più ricche dei modelli di $\varphi''$, ma che *algebricamente* saranno del loro stesso tipo. Ecco perché, ad esempio, sia (m2) che la sua skolemiana (m2') sono assiomi per la teoria dei monoidi, cossiccome (g) e la sua skolemiana (g') lo sono della teoria dei gruppi, e (k2) e (k2') lo sono dei campi. Tuttavia in questi esempi da noi considerati l'utilizzo di forme skolemizzate risultava del tutto naturale, in quanto gli elementi neutri sono unici e quelli inversi univocamente determinati per ogni argomento dato. Qualora tali determinazioni precise vengano a cadere, l'utilizzo delle skolemizzazioni può apparire meno naturale, ed in effetti è poco usato nella normale pratica matematica.

**Esempio 5.24.** *Consideriamo l'assioma di densità valido sulla struttura dei numeri razionali* $\mathbb{Q}$

$$(\forall x_1)(\forall x_2)(\exists y)(x_1 < x_2 \to x_1 < y < x_2).$$

*Questo assioma si può skolemizzare come*

$$(\forall x_1)(\forall x_2)(x_1 < x_2 \to x_1 < f(x_1, x_2) < x_2)$$

*dove* $f(x_1, x_2)$ *è uno dei tanti (di fatto infiniti!) elementi che giacciono tra* $x_1$ *ed* $x_2$*. Tuttavia poiché l'elemento intermedio non è univocamente determinato, l'utilizzo della versione skolemizzata dell'assioma di densità non risulta essere molto naturale. Si noti però che in molti casi concreti la scelta di un valore preciso tra i tanti possibili può essere resa più raffinata fino a diventare computabile. Ad esempio in* $\mathbb{Q}$ *potremmo porre* $f(x_1, x_2) := x_1 + \frac{x_2 - x_1}{2}$.

Data una formula $\varphi$ chiusa sia $\varphi^S$ una sua forma skolemizzata lungo le linee della dimostrazione del Teorema di Skolem 5.22. Allora $\varphi^S$ è della forma $(\forall x_1)...(\forall x_n)\,\varphi_S(x_1,...,x_n)$, dove $\varphi_S(x_1,...,x_n)$ non contentente quantificatori è detta *matrice skolemiana* di $\varphi$.

Possiamo anche skolemizzare un insieme di $\mathscr{L}$-formule $\Gamma$ ottendendo un insieme $\Gamma^S$ di forme skolemiane, in un linguaggio $\mathscr{L}'$, delle formule di $\Gamma$ in modo che per ogni $\varphi \in \Gamma$ i nuovi simboli di funzione introdotti per $\varphi^S$ non siano stati usati per le altre formule di $\Gamma$. Sia ora $\Gamma^g$ l'insieme delle istanze *ground* di $\Gamma^S$, ovvero l'insieme di tutte le formule del tipo $\varphi_S(t_1/x_1,..,t_n/x_n)$ per $\varphi \in \Gamma$ e $t_1,...,t_n$ $\mathscr{L}'$-termini chiusi (se $\mathscr{L}$ non contiene costanti e la skolemizzazione non ci porta ad introdurre nuove costanti, ne introduciamo una *ex novo*).

**Definizione 5.25.** *Dato un insieme* $\Delta$ *di* $\mathscr{L}$*-formule chiuse senza quantificatori, sia* $\Delta_{atom}$ *l'insieme di tutte le sottoformule atomiche di formule di* $\Delta$.

*Dato un qualsiasi assegmanento di valori di verità* $\sigma : \Delta_{atom} \to \{0,1\}$*, estendiamo tale* $\sigma$ *a tutto* $\Delta$ *in modo che per ogni* $\varphi \in \Delta$*,* $\hat{\sigma}(\varphi)$ *sia il valore di verità assegnato a* $\varphi$ *dalla sua tavola di verità sulla base dei valori assegnati da* $\sigma$ *alle sue formule atomiche.*

*Allora diciamo che* $\Delta$ *è* proposizionalmente soddisfacibile *se esiste un assegmanento di valori di verità* $\sigma : \Delta_{atom} \to \{0,1\}$ *tale che* $\hat{\sigma}(\psi) = 1$ *per ogni* $\psi \in \Delta$.

L'idea di questa definizione è quella di considerare le formule ground senza quantificatori come formule proposizionali.

Il seguente teorema riduce quindi la soddisfacibilità predicativa a quella proposizionale.

**Teorema 5.26** (Teorema di Herbrand)**.** *Un insieme* $\Gamma$ *di* $\mathscr{L}$*-formule chiuse senza identità è soddisfacibile sse* $\Gamma^g$ *è proposizionalmente soddisfacibile.*

*Dimostrazione.* $\Rightarrow$) $\Gamma$ soddisfacibile $\Rightarrow$ $\Gamma^S$ soddisfacibile (per estensione naturale del Teorema di Skolem 5.22) $\Rightarrow$ $\Gamma^g$ soddisfacibile (per Ax. Q1 si ha che $(\forall x_1)...(\forall x_n)\varphi_S(x_1,...,x_n) \models \varphi_S(t_1/x_1,...,t_n/x_n)$ per arbitrari $\mathscr{L}$-termini chiusi $t_1,...,t_n$) $\Rightarrow$ $\Gamma^g$ è proposizionalmente soddisfacibile (basta porre $\sigma(P(t_1/x_1,...,t_n/x_n)) := 1$ se e solo se $\mathscr{M} \models P(t_1/x_1,...,t_n/x_n)$ per $\mathscr{M}$ modello di $\Gamma^g$).

$\Leftarrow$) Supponiamo che $\Gamma^g$ sia proposizionalmente soddisfacibile, e sia quindi data $\sigma : \Gamma^g_{atom} \to \{0,1\}$ tale che $\hat{\sigma}(\varphi) = 1$ per ogni $\varphi \in \Gamma^g$. Definiamo allora un modello di Herbrand $\mathscr{M}$ per $\Gamma^g$ in modo analogo alla costruzione dei modelli di Henkin, per $\mathscr{L}'$ il linguaggio di $\Gamma^g$:

- $M$ è l'insieme dei termini chiusi di $\mathscr{L}'$;
- $c^{\mathscr{M}} := c$ per ogni costante $c$ di $\mathscr{L}'$;
- $(f^n)^{\mathscr{M}} : M^n \to M, (t_1,..,t_n) \mapsto f^n(t_1,...,t_n)$,
  cioè $(f^n)^{\mathscr{M}}(t_1,..,t_n) := f^n(t_1,..,t_n)$ per ogni $\mathscr{L}'$-simbolo di funzione $f$;
- $(t_1,...,t_n) \in (P^n)^{\mathscr{M}} \subseteq M^n$ sse $\sigma(P(t_1,..,t_n)) = 1$ per ogni $\mathscr{L}'$-simbolo di predicato $P$.

È immediato dimostrare, per induzione, che per ogni sottoformula $\psi$ di qualche formula in $\Gamma^g$ (dunque $\psi$ è chiusa e senza quantificatori) si ha $\mathscr{M} \models \psi \Leftrightarrow \hat{\sigma}(\psi) = 1$:

- $\psi \equiv P(t_1, ..., t_n)$: segue immediatamente dalla definizione di modello di Herbrand
- $\psi \equiv \neg\chi$: $\mathscr{M} \models \neg\chi$ sse $\mathscr{M} \not\models \chi$ sse $\hat{\sigma}(\chi) = 0$ (IH) sse $\hat{\sigma}(\neg\chi) = 1$
- $\psi \equiv \chi \wedge \xi$: $\mathscr{M} \models \chi \wedge \xi$ sse $\mathscr{M} \models \chi$ e $\mathscr{M} \models \xi$ sse $\hat{\sigma}(\chi) = 1$ e $\hat{\sigma}(\xi) = 1$ (IH) sse $\hat{\sigma}(\chi \wedge \xi) = 1$
- $\psi \equiv \chi \vee \xi, \chi \to \xi$: analogo

Pertanto $\mathscr{M} \models \Gamma^g$.

Dimostriamo ora che anche $\mathscr{M} \models \Gamma^S$. Sia

$$\varphi^S \equiv (\forall x_1)...(\forall x_n)\varphi_S(x_1, ..., x_n) \in \Gamma^S.$$

Per ogni $t_1, ..., t_n \in M$, $\varphi_S(t_1/x_1, ..., t_n/x_n) \in \Gamma^g$.

Quindi $\hat{\sigma}(\varphi_S(t_1/x_1, ..., t_n/x_n)) = 1$, da cui, per quanto appena visto, $\mathscr{M} \models \varphi_S(t_1/x_1, ..., t_n/x_n)$. Poiché $t_1, .., t_n$ sono generici elementi di $M$, abbiamo che $\mathscr{M} \models (\forall x_1)...(\forall x_n)\varphi_S(x_1, ..., x_n)$. □

**Esempio 5.27.**

- *Verifichiamo se* $(\exists x)P(x), (\forall x)(P(x) \to Q(x)) \models (\exists x)Q(x)$.

  *Osserviamo che* $(\exists x)P(x), (\forall x)(P(x) \to Q(x)) \models (\exists x)Q(x)$

  *sse* $\{(\exists x)P(x), (\forall x)(P(x) \to Q(x)), \neg(\exists x)Q(x)\}$ *insodd. sse* $\Gamma := \{(\exists x)P(x), (\forall x)(P(x) \to Q(x)), (\forall x)\neg Q(x)\}$ *insodd. sse*

  $$\Gamma^S := \{P(c/x), (\forall x)(P(x) \to Q(x)), (\forall x)\neg Q(x)\}$$

  *è insodd. sse* $\Gamma^g$ *prop. insodd. Ma* $\Gamma^g$ *è proposizionalmente insoddisfacibile, in quanto* $\{P(c/x), P(c/x) \to Q(c/x), \neg Q(c/x)\} \subseteq \Gamma^g$ *è proposizionalmente insoddisfacibile.*

- *Verifichiamo se* $(\forall x)(\exists y)P(x, y) \models (\exists y)(\forall x)P(x, y)$.

  *Osserviamo che* $(\forall x)(\exists y)P(x, y) \models (\exists y)(\forall x)P(x, y)$

  *sse* $\{(\forall x)(\exists y)P(x, y), \neg(\exists y)(\forall x)P(x, y)\}$ *insodd.*

*sse* $\Gamma := \{(\forall x)(\exists y)P(x,y), (\forall y)(\exists x)\neg P(x,y)\}$ *insodd. sse*

$$\Gamma^S := \{(\forall x)P(x, f(x)), (\forall y)\neg P(g(y), y)\}$$

*è insodd sse* $\Gamma^g$ *prop. insodd. Ma* $\Gamma^g$ *è proposizionalmente soddisfacibile: basta porre* $\sigma(P(t_1, f(t_1)) := 1$ *e* $\sigma(P(g(t_2), t_2) := 0$ *per tutti gli* $\mathscr{L}$*-termini chiusi* $t_1, t_2$.

- *Verifichiamo se* $\models (\exists y)(\forall x)(P(y) \to P(x))$.

  *Osserviamo che* $\models (\exists y)(\forall x)(P(y) \to P(x))$ *sse* $\{\neg(\exists y)(\forall x)(P(y) \to P(x))\}$ *insodd sse* $\Gamma := \{(\forall y)(\exists x)(P(y) \wedge \neg P(x))\}$ *insodd sse* $\Gamma^S := \{(\forall y)(P(y) \wedge \neg P(f(y)))\}$ *insodd sse* $\Gamma^g$ *prop. insodd. Ma* $\Gamma^g$ *è proposizionalmente insoddisfacibile in quanto l'insieme*

  $$\{P(c) \wedge \neg P(f(c)), P(f(c)) \wedge \neg P(f(f(c)))\} \subseteq \Gamma^g$$

  *è proposizionalmente insoddisfacibile.*

- *Verifichiamo se* $(\forall x)(P(x) \vee Q(x)) \models (\forall x)P(x) \vee (\forall y)Q(y)$.

  *Osserviamo che* $(\forall x)(P(x) \vee Q(x)) \models (\forall x)P(x) \vee (\forall y)Q(y)$ *sse* $\{(\forall x)(P(x) \vee Q(x)), \neg[(\forall x)P(x) \vee (\forall y)Q(y)]\}$ *insodd.* *sse* $\Gamma := \{(\forall x)(P(x) \vee Q(x)), (\exists x)\neg P(x) \wedge (\exists y)\neg Q(y)\}$ *insodd. sse* $\Gamma^S := \{(\forall x)(P(x) \vee Q(x)), \neg P(c) \wedge \neg Q(d)\}$ *insodd. sse* $\Gamma^g$ *prop. insodd. Ma*

  $$\Gamma^g := \big\{P(t) \vee Q(t) \mid t\ \mathscr{L}\text{-termine chiuso}\big\} \cup \{\neg P(c) \wedge \neg Q(d)\}$$

  *è proposizionalmente soddisfacibile: basta porre* $\sigma(P(c)) := 0$, $\sigma(Q(d)) := 0$, $\sigma(Q(c)) := 1$, $\sigma(P(t)) := 1$ *per* $t \not\equiv c$.

**Corollario 5.28.** *Sia data una* $\mathscr{L}$*-formula* $(\exists x)\psi$ *con* $\psi$ *senza quantificatori e senza identità e tale che* $\vdash (\exists x)\psi$. *Allora esistono* $\mathscr{L}$*-termini chiusi* $t_1, \ldots, t_n$ *per un certo* $n \geq 1$ *tali che* $\vdash \psi(t_1) \vee \ldots \vee \psi(t_n)$.

*Dimostrazione.* Supponiamo $\nvdash \psi(t_1) \vee \ldots \vee \psi(t_n)$ per tutti gli $n \geq 1$ e tutti gli $\mathscr{L}$-termini chiusi $t_1, \ldots, t_n$; allora $\not\models \psi(t_1) \vee \ldots \vee \psi(t_n)$ per tutti gli $n \geq 1$ e tutti gli $\mathscr{L}$-termini chiusi $t_1, \ldots, t_n$ (per Teorema di Completezza), da cui $\neg(\psi(t_1) \vee \ldots \vee \psi(t_n))$ è soddisfacibile e quindi $\neg\psi(t_1) \wedge \ldots \wedge \neg\psi(t_n)$ è soddisfacibile. Perciò per ogni $n \geq 1$ e $t_1, \ldots, t_n$

$\mathscr{L}$-termini chiusi si ha che $\{\neg\psi(t_1), \ldots, \neg\psi(t_n)\}$ è soddisfacibile. Per il Teorema di Compattezza (per ogni cardinalità)

$$\{\neg\psi(t) \mid t \text{ è } \mathscr{L}\text{-termine chiuso}\}$$

è soddisfacibile, e quindi, essendo le $\neg\psi(t)$ chiuse senza quantificatori, anche proposizionalmente soddisfacibile. Per il Teorema di Herbrand 5.26, $(\forall x)\,\neg\psi(x)$ è soddisfacibile (dove $\Gamma := \{(\forall x)\,\neg\psi(x)\} = \Gamma^S$), da cui $\neg(\exists x)\,\psi(x)$ è soddisfacibile, e quindi, per Teorema di Validità $\not\models (\exists x)\,\psi(x)$. $\square$

# Capitolo 6

# Categoricità

**Definizione 6.1** (Teorie categoriche)**.** *Sia $k$ un cardinale e sia $T$ una $\mathscr{L}$-teoria con $\mathscr{L}$-modelli di cardinalità $k$. Allora $T$ si dice $k$-*categorica *se tutti gli $\mathscr{L}$-modelli di $T$ di cardinalità $k$ sono tra loro isomorfi.*

Un altro modo di esprimere la questione è dire che $T$ ha un unico modello di cardinalità $k$ a meno di isomorfismi.

Consideriamo la teoria degli ordini lineari densi stretti senza massimi né minimi nel linguaggio $\mathscr{L} := \{<\}$. Questa teoria è caratterizzata dai seguenti assiomi:

- $(\forall x)\, \neg x < x$ (irriflessività)
- $(\forall x)(\forall y)(\forall z)(x < y \rightarrow (y < z \rightarrow x < z))$ (transitività)
- $(\forall x)(\forall y)(x < y \vee x = y \vee y < x)$ (linearità)
- $(\forall x)(\forall y)(\exists z)(x < y \rightarrow x < z < y)$ (densità)
- $(\forall x)(\exists y)(\exists z)\, y < x < z$ (illimitatezza)

**Teorema 6.2.** *La teoria $T$ degli ordini lineari densi stretti senza massimi né minimi è $\aleph_0$-categorica.*

*Dimostrazione.* Ovviamente $T$ ha modelli infiniti numerabili (ad esempio $\mathbb{Q}$).

Siano ora $\mathscr{A} := \langle A, < \rangle$ e $\mathscr{B} := \langle B, < \rangle$ due ordini lineari densi stretti numerabili senza massimi né minimi, e siano $a_0, a_1, a_2, \ldots, a_n, \ldots$ e $b_0, b_1, b_2, \ldots, b_n, \ldots$ due enumerazioni (senza ripetizioni) dei domini $A$ e $B$ rispettivamente. Vediamo come costruire un isomorfismo $h : A \to B$ che preservi $<$ mediante la tecnica induttiva del *back and forth*:

passo $p := 0$) poniamo $h(a_0) := b_0$;

passo $s := 2m+1$) siano già stati definiti ai passi precedenti $h(a_{i_0}) < \ldots < h(a_{i_k}) \in B$ per $a_{i_0} < \ldots < a_{i_k} \in A$. Consideriamo ora $a_m \in A$. Se $a_m = a_{i_j}$ per un $j \in \{0 \ldots, k\}$, allora passiamo al passo successivo. Altrimenti poiché l'ordine su $A$ è lineare si darà uno ed uno solo dei seguenti casi:

(i) $a_{i_j} < a_m$ per $0 \leq j \leq k$. Poiché in $B$ non vi è un massimo, esiste un $b$ tale che $h(a_{i_j}) < b$ per $0 \leq j \leq k$. Poniamo allora $h(a_m) := b$;

(ii) $a_m < a_{i_j}$ per $0 \leq j \leq k$. Poiché in $B$ non vi è un minimo, esiste un $b$ tale che $b < h(a_{i_j})$ per $0 \leq j \leq k$. Poniamo allora $h(a_m) := b$;

(iii) $a_{i_j} < a_m < a_{i_{j+1}}$ per (esattamente) un $j \in \{0 \ldots, k\}$. Poiché l'ordine in $B$ è denso, esiste un $b$ tale che $h(a_{i_j}) < b < h(a_{i_{j+1}})$. Poniamo allora $h(a_m) := b$.

passo $s := 2m+2$) siano già stati definiti ai passi precedenti $h(a_{i_0}) < \ldots < h(a_{i_k}) \in B$ per $a_{i_0} < \ldots < a_{i_k} \in A$. Consideriamo ora $b_m \in B$. Se $b_m = h(a_{i_j})$ per qualche $j \in \{0 \ldots, k\}$, allora passiamo al passo successivo. Altrimenti poiché l'ordine su $B$ è lineare si darà uno ed uno solo dei seguenti casi:

(i) $h(a_{i_j}) < b_m$ per $0 \leq j \leq k$. Poiché in $A$ non vi è un massimo, esiste un $a$ tale che $a_{i_j} < a$ per $0 \leq j \leq k$. Poniamo allora $h(a) := b_m$;

(ii) $b_m < h(a_{i_j})$ per $0 \leq j \leq k$. Poiché in $A$ non vi è un minimo, esiste un $a$ tale che $a < a_{i_j}$ per $0 \leq j \leq k$. Poniamo allora $h(a) := b_m$;

(iii) $h(a_{i_j}) < b_m < h(a_{i_{j+1}})$ per (esattamente) un $j \in \{0 \ldots, k\}$. Poiché l'ordine in $A$ è denso, esiste un $a$ tale che $a_{i_j} < a < a_{i_{j+1}}$. Poniamo allora $h(a) := b_m$.

Grazie ai passi dispari, la funzione $h$ è definita per tutto $A$. Inoltre $h$ rispetta l'ordine per costruzione (esercizio). Da questo segue anche che $h$ è iniettiva: siano infatti $a \neq a'$ per $a, a' \in A$; allora per linearità, $a < a'$ o $a' < a$. Si dia ad esempio il primo caso (il secondo è analogo). Allora $h(a) < h(a')$, da cui $h(a) \neq h(a')$ per irriflessività. Infine, grazie ai passi pari $h$ è suriettiva. Pertanto $h$ è un isomorfismo tra $\mathscr{A}$ e $\mathscr{B}$. $\square$

Tuttavia la proprietà non è preservata oltre $\aleph_0$. Vediamo ad esempio che $T$ non è $2^{\aleph_0}$-categorica. Consideriamo $\mathbb{R}$ ed $\mathbb{R} \setminus \{0\} = \mathbb{R}^- \cup \mathbb{R}^+$. Ovviamente sia $\mathbb{R}$ che $\mathbb{R}^- \cup \mathbb{R}^+$ sono ordini lineari densi stretti senza massimo né minimo di cardinalità $2^{\aleph_0}$. Vediamo tuttavia che $\mathbb{R}$ e $\mathbb{R}^- \cup \mathbb{R}^+$ non sono strutture isomorfe. Supponiamo per assurdo che lo siano tramite l'isomorfismo $h : \mathbb{R}^- \cup \mathbb{R}^+ \to \mathbb{R}$. Consideriamo ora $h(\mathbb{R}^-) \subseteq \mathbb{R}$. In quanto isomorfismo, $h$ deve preservare l'ordine, per cui $h(-r) < h(s)$ per ogni $-r \in \mathbb{R}^-$, $s \in \mathbb{R}^+$. Pertanto $h(\mathbb{R}^-)$ ha un estremante superiore (maggiorante[1]) in $\mathbb{R}$. Ma $\mathbb{R}$ è *completo*, ovvero ogni insieme non vuoto che abbia un maggiorante (risp. minorante), ha un supremo (risp. infimo), e quindi $h(\mathbb{R}^-)$ avrà un supremo $\sigma$. Per suriettività di $h$ si ha che $\sigma \in h(\mathbb{R}^-)$ o $\sigma \in h(\mathbb{R}^+)$. Se $\sigma \in h(\mathbb{R}^-)$ allora esiste un $-r \in \mathbb{R}^-$ tale che $h(-r) = \sigma$. Poiché $-r < -r + \frac{r}{2} < 0$, si ha $\sigma = h(-r) < h(-r + \frac{r}{2}) \in h(\mathbb{R}^-)$, il che è assurdo. Se invece $\sigma \in h(\mathbb{R}^+)$, allora esiste un $r \in \mathbb{R}^+$ tale che $h(r) = \sigma$. Poiché $0 < r - \frac{r}{2} < r$, si ha $h(r - \frac{r}{2}) < h(r)$. Ma $r - \frac{r}{2} \in \mathbb{R}^+$, e quindi per ogni $-t \in \mathbb{R}^-$ avremo $h(-t) < h(r - \frac{r}{2}) < h(r) = \sigma$, il che è assurdo.

Consideriamo ora la teoria $T$ dei gruppi abeliani senza torsioni nel linguaggio $\mathscr{L} := \{\dot{0}, \dot{+}\}$. Questa teoria è caratterizzata dagli assiomi di gruppo abeliano in $\mathscr{L}$ ed inoltre:

$$\text{(gast1)} \quad (\forall x)(x \neq \dot{0} \to \underbrace{x \dot{+} x \dot{+} \dots \dot{+} x}_{n\text{-volte}} \neq \dot{0})$$

$$\text{(gast2)} \quad (\forall x)(\exists y)\, \underbrace{y \dot{+} y \dot{+} \dots \dot{+} y}_{n\text{-volte}} = x$$

per $n \geq 1$ (si osservi quindi che gast1 e gast2 sono schemi di assiomi).

[1]Dato $A \subseteq X$ e $b \in X$, per $X$ insieme (parzialmente) ordinato, diciamo che $b$ è *maggiorante* (*estremante superiore*) di $A$ se $a \leq b$ per ogni $a \in A$. L'elemento $b$ è anche *supremo* di $A$ se è il più piccolo dei suoi maggioranti. I concetti di *minorante* (*estremante inferiore*) ed *infimo* sono definiti in modo duale.

**Teorema 6.3.** *La teoria $T$ dei gruppi abeliani senza torsione è $k$-categorica, per $k$ transfinito, sse $k > \aleph_0$.*

*Dimostrazione.* Dimostriamo che i modelli di $T$ sono esattamente i gruppi abeliani che costituiscono uno spazio vettoriale sul campo $\mathbb{Q}$.

Da un lato è facile vedere che un gruppo $\mathscr{G} := \langle G, \dot{0}, \dot{+} \rangle$ di questo tipo soddisfa, oltre ovviamente gli assiomi dei gruppi abeliani, anche i due assiomi specifici di questa teoria.

Per quanto riguarda (gast1), osserviamo che

$$\underbrace{x \dot{+} \ldots \dot{+} x}_{n\text{-volte}} = \underbrace{1x \dot{+} \ldots \dot{+} 1x}_{n\text{-volte}} = (\underbrace{1 + \ldots + 1}_{n\text{-volte}})x = nx$$

per (v1) e (v4). Supponiamo che sia $nx = \dot{0}$. Giacché $n \neq 0$, per la Proposizione 3.11.3, abbiamo che $x = \dot{0}$.

Per quanto riguarda (gast2), osserviamo che

$$\underbrace{\frac{1}{n}x \dot{+} \ldots \dot{+} \frac{1}{n}x}_{n\text{-volte}} = (\underbrace{\frac{1}{n} \dot{+} \ldots \dot{+} \frac{1}{n}}_{n\text{-volte}})x = 1x = x$$

ancora per (v1) e (v4).

Sia ora invece $\mathscr{G} \models T$ per $G$ dominio di $\mathscr{G}$. Ovviamente $\mathscr{G}$ è un gruppo abeliano. Mostriamo poi che $\mathscr{G}$ può essere considerato uno spazio vettoriale sul campo $\mathbb{Q}$. Sia infatti $g \in G$. Allora per $n \geq 1$ sappiamo per (gast2) che esiste un $h$ tale che $\underbrace{h \dot{+} h \dot{+} \ldots \dot{+} h}_{n\text{-volte}} = g$. Mostriamo anche che tale $h$ è univocamente determinato. Soddisfi infatti $h'$ la stessa proprietà. Allora $\dot{0} = g \dot{-} g = \underbrace{h \dot{+} h \dot{+} \ldots \dot{+} h}_{n\text{-volte}} - (\underbrace{h' \dot{+} h' \dot{+} \ldots \dot{+} h'}_{n\text{-volte}})$.
Per la Proposizione 3.1.1 (che di fatto si applica già ai gruppi abeliani), sappiamo che l'ultimo termine è uguale a

$$\underbrace{h \dot{+} h \dot{+} \ldots \dot{+} h}_{n\text{-volte}} \underbrace{\dot{-} h' \dot{-} h' \dot{-} \ldots \dot{-} h'}_{n\text{-volte}},$$

da cui per commutatività ed associatività segue

$$\dot{0} = \underbrace{(h \dot{-} h') \dot{+} \ldots \dot{+} (h \dot{-} h')}_{n\text{-volte}}.$$

Poichè $\mathscr{G}$ è senza torsioni per (gast1), questo si dà solo se $h \dot{-} h' = \dot{0}$, cioè $h = h'$. Denotiamo quest'unico $h$ come $\frac{g}{n}$. Possiamo quindi vedere che $\mathscr{G}$ può essere inteso come uno spazio vettoriale su $\mathbb{Q}$ mediante la traduzione

- $mg := \underbrace{g \dot{+} \dots \dot{+} g}_{m\text{-volte}}$
- $0g := \dot{0}$
- $\frac{m}{n} g := m \frac{g}{n}$
- $-\frac{m}{n} g := \frac{m}{n} (\dot{-} g)$

per $g \in G$, $m, n \in \mathbb{N} \setminus \{0\}$.

Vediamo infatti che gli assiomi degli spazi vettoriali sono soddisfatti:

(v1) $1 \cdot g = g$: si dà infatti $1 \cdot g = \underbrace{g}_{1\text{-volta}} = g$.

(v2) $\frac{m'}{n'}(\frac{m}{n} g) = (\frac{m'}{n'} \frac{m}{n}) g$: mostriamo che $\frac{m'}{n'}(\frac{m}{n} g)$ coincide con il vettore $\frac{m'm}{n'n} g$, ricordando poi che su $\mathbb{Q}$ il termine $\frac{m'm}{n'n}$ corrisponde a $\frac{m'}{n'} \frac{m}{n}$. Notiamo innanzitutto che $k(k'h) = (kk')h = (k'k)h = k'(kh)$ per $k, k' \in \mathbb{N}$ ed $h \in G$ (per definizione). Mostriamo ora che per ogni $h \in G$ e $k, k' \in \mathbb{N} \setminus \{0\}$ si ha $\frac{\frac{h}{k}}{k'} = \frac{h}{k'k}$; vediamo cioè che $\frac{\frac{h}{k}}{k'}$ è la $(k'k)$-esima radice di $h$ (la quale è univocamente determinata). Infatti

$$\underbrace{\frac{\frac{h}{k}}{k'} \dot{+} \dots \dot{+} \frac{\frac{h}{k}}{k'}}_{k'k\text{-volte}} = \underbrace{\underbrace{\frac{\frac{h}{k}}{k'} \dot{+} \dots \dot{+} \frac{\frac{h}{k}}{k'}}_{k'\text{-volte} = \frac{h}{k}} \dot{+} \dots \dot{+} \underbrace{\frac{\frac{h}{k}}{k'} \dot{+} \dots \dot{+} \frac{\frac{h}{k}}{k'}}_{k'\text{-volte} = \frac{h}{k}}}_{k\text{-volte} = h}.$$

Ora vediamo che $k \frac{h}{k'} = \frac{kh}{k'}$ per ogni $h \in G$ e $k' \in \mathbb{N} \setminus \{0\}$, vediamo cioè che $k \frac{h}{k'}$ è la $k'$-esima radice di $kh$ (la quale è univocamente determinata):

$$\underbrace{\underbrace{\frac{h}{k'}\dot{+}\dots\dot{+}\frac{h}{k'}}_{k\text{-volte}}\dot{+}\dots\dot{+}\underbrace{\frac{h}{k'}\dot{+}\dots\dot{+}\frac{h}{k'}}_{k\text{-volte}}}_{k'\text{-volte}} = \underbrace{\frac{h}{k'}\dot{+}\dots\dot{+}\frac{h}{k'}}_{k'k\text{-volte}} =$$

$$= \underbrace{\underbrace{\frac{h}{k'}\dot{+}\dots\dot{+}\frac{h}{k'}}_{k'\text{-volte}=h}\dot{+}\dots\dot{+}\underbrace{\frac{h}{k'}\dot{+}\dots\dot{+}\frac{h}{k'}}_{k'\text{-volte}=h}}_{k\text{-volte } = kh}.$$

Concludiamo che $\frac{m'}{n'}(\frac{m}{n}g) = \frac{m'}{n'}(m\frac{g}{n}) = m'\frac{m\frac{g}{n}}{n'} = m'\left(m\frac{\frac{g}{n}}{n'}\right) = (m'm)\frac{\frac{g}{n}}{n'} = (m'm)\frac{g}{n'n} = \frac{m'm}{n'n}g = (\frac{m'}{n'}\frac{m}{n})g.$

(v3) $\frac{m}{n}(g\dot{+}g') = \frac{m}{n}g\dot{+}\frac{m}{n}g'$: per prima cosa osserviamo che per commutatività ed associatività è facile vedere che $k(h\dot{+}h') = kh\dot{+}kh'$ per $h, h' \in G$ e $k \in \mathbb{N}$. Vediamo quindi che $\frac{g}{n}\dot{+}\frac{g'}{n} = \frac{g\dot{+}g'}{n}$, cioè $\frac{g}{n}\dot{+}\frac{g'}{n}$ è l'$n$-esima radice di $g\dot{+}g'$: $\underbrace{(\frac{g}{n}\dot{+}\frac{g'}{n})\dot{+}\dots\dot{+}(\frac{g}{n}\dot{+}\frac{g'}{n})}_{n\text{-volte}} =$

$$\underbrace{\frac{g}{n}\dot{+}\dots\dot{+}\frac{g}{n}}_{n\text{-volte } = g}\dot{+}\underbrace{\frac{g'}{n}\dot{+}\dots\dot{+}\frac{g'}{n}}_{n\text{-volte } = g'} = g\dot{+}g'.$$

Concludiamo che $\frac{m}{n}(g\dot{+}g') = m\frac{g\dot{+}g'}{n} = m(\frac{g}{n}\dot{+}\frac{g'}{n}) = m\frac{g}{n}\dot{+}m\frac{g'}{n} = \frac{m}{n}g\dot{+}\frac{m}{n}g'$.

(v4) $(\frac{m'}{n'}+\frac{m}{n})g = \frac{m'}{n'}g\dot{+}\frac{m}{n}g$: mostriamo che $(\frac{m'}{n'}+\frac{m}{n})g$ coincide con il vettore $\frac{m'}{n'}g\dot{+}\frac{m}{n}g$ leggendo $\frac{m'}{n'}+\frac{m}{n}$, come avviene in $\mathbb{Q}$, come $\frac{m'n+mn'}{n'n}$. È immediato vedere che $(k+k')h = kh\dot{+}k'h$ per $h \in G$ e $k, k' \in \mathbb{N}$. Sulla base di quanto dimostrato nei casi precedenti possiamo allora dedurre che $(\frac{m'}{n'}+\frac{m}{n})g = \frac{m'n+mn'}{n'n}g = (m'n+mn')\frac{g}{n'n} = (m'n)\frac{g}{n'n}\dot{+}(mn')\frac{g}{n'n} = (m'n)\frac{g}{nn'}\dot{+}(mn')\frac{g}{n'n} = m'n\frac{\frac{g}{n'}}{n}\dot{+}mn'\frac{\frac{g}{n}}{n'} = m'\frac{g}{n'}\dot{+}m\frac{g}{n} = \frac{m'}{n'}g\dot{+}\frac{m}{n}g.$

Pertanto i modelli di $T$ sono esattamente i gruppi che sono spazi vettoriali su $\mathbb{Q}$. Da questo segue innanzitutto che $T$ ha modelli di qualsiasi cardinalità $k \geq \aleph_0$. Vediamo ora che $T$ è $k$-categorica per $k > \aleph_0$. Consideriamo infatti due gruppi $\mathscr{G}$ e $\mathscr{G}'$ che siano spazi vettoriali su $\mathbb{Q}$ e che abbiano cardinalità $k > \aleph_0$. Allora $\mathscr{G}$ e $\mathscr{G}'$ come spazi vettoriali hanno basi di cardinalità $k$ (si riconsideri l'Esempio 3.17.3), e quindi hanno dimensione $k$. Per il Teorema della Dimensione 3.20 $\mathscr{G}$ e $\mathscr{G}'$ sono isomorfi in quanto spazi vettoriali. In particolare sono isomorfi anche come gruppi sul linguaggio $\mathscr{L}$: l'asserzione è ovvia per $\dot{+}$, mentre per $\dot{0}$ si ricava dalla Proposizione 3.19.2.

Vediamo che tuttavia $T$ non è $\aleph_0$-categorica. Consideriamo gli spazi vettoriali $\mathbb{Q}, \mathbb{Q}^2$. Entrambi hanno $\aleph_0$-elementi. Vediamo ora che tali spazi non sono isomorfi come gruppi abeliani nel linguaggio $\mathscr{L}$. Non è infatti possibile definire alcuna biezione $h : \mathbb{Q} \to \mathbb{Q}^2$ che preservi $+$ e $0$ tra tali gruppi abeliani (tale cioè che $h(0) = 0$ e $h(x+y) = h(x) + h(y)$)[2]. Supponiamo infatti per assurdo che ciò si dia. Allora è facile vedere che $h$ preserverebbe anche il prodotto scalare per tutti i coefficienti del campo $\mathbb{Q}$. Infatti per $\mathbb{N}^+$ la proprietà seguirebbe dagli assiomi (v1) e (v4) della Definizione 3.10 e per 0 seguirebbe dalla Proposizione 3.11.1 (esercizio). Per $\mathbb{Z}$ osserviamo che avremmo $0 = h(0) = h(x-x) = h(x) + h(-x)$, da cui $h(-x) = -h(x)$. Infine per l'intero $\mathbb{Q}$ osserviamo che $h(\frac{p}{q}x) = h(p\frac{x}{q}) = ph(\frac{x}{q}) = p\frac{q}{q}h(\frac{x}{q}) = \frac{p}{q}qh(\frac{x}{q}) = \frac{p}{q}h(q\frac{x}{q}) = \frac{p}{q}h(x)$. Ma allora $h$ sarebbe un isomorfismo tra spazi vettoriali, il che contraddice il Teorema della Dimensione 3.20. □

Abbiamo quindi visto un esempio di teoria categorica solo su $\aleph_0$, ed una, all'opposto, che è categorica solo per transfiniti $k > \aleph_0$. Vediamo ora un esempio di teoria categorica per *ogni* cardinale transfinito. Si consideri al riguardo la teoria dei gruppi abeliani ad elementi di ordine 2, caratterizzata dagli assiomi dei gruppi abeliani nel linguaggio $\mathscr{L} := \{\dot{0}, \dot{+}\}$ più l'assioma

(ord2) $(\forall x)\, x \dot{+} x = \dot{0}$

[2] Per semplicità notazionale, trasferiamo la notazione di $\mathbb{Q}$ su $\mathbb{Q}^2$. In $\mathbb{Q}^2$ la somma è definita come in Esempio 3.17.2, mentre il vettore 0 è il vettore $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$; dato $x \in \mathbb{Q}^2$, l'elemento $-x \in \mathbb{Q}^2$ è definito di conseguenza.

**Teorema 6.4.** *La teoria $T$ dei gruppi abeliani ad elementi di ordine 2 è $k$-categorica per ogni cardinale transfinito $k$.*

*Dimostrazione.* In maniera analoga alla dimostrazione del Teorema 6.3 mostriamo che i modelli di $T$ sono esattamente i gruppi abeliani che possiamo vedere come spazi vettoriali su $\mathbb{Z}_2$. Da un lato è immediato vedere che ogni gruppo abeliano $\mathcal{G} := \langle G, \dot{0}, \dot{+} \rangle$ di questo tipo soddisfa gli assiomi dei gruppi abeliani più (ord2); per (ord2) si osservi semplicemente che $x \dot{+} x = 1 \cdot x \dot{+} 1 \cdot x = (1 +_2 1)x = 0 \cdot x = \dot{0}$ (dove $+_2$ è la somma in $\mathbb{Z}_2$) per (v1), (v4) e Proposizione 3.11.1.

Per l'altra direzione, sia $\mathcal{G} \models T$ per $G$ dominio di $\mathcal{G}$. Chiaramente $\mathcal{G}$ è un gruppo abeliano. Definiamo inoltre

- $0 \cdot g = \dot{0}$.
- $1 \cdot g = g$.

Vediamo che tale prodotto soddisfa gli assiomi (v1)-(v4) degli spazi vettoriali (analogamente a $+_2$, sia $\cdot_2$ l'operazione di prodotto su $\mathbb{Z}_2$):

(v1) $1 \cdot g = g$: per definizione

(v2) $m(ng) = (m \cdot_2 n)g$: distinguiamo tra due casi. Per $n = 0$ si ha $m(0g) = m\dot{0} = \dot{0}$ (esercizio: per definizione di $\cdot$) $= 0g = (m \cdot_2 0)g$. Per $n = 1$ si ha $m(1g) = mg = (m \cdot_2 1)g$.

(v3) $n(g \dot{+} g') = ng \dot{+} ng$: distinguiamo tra due casi. Per $n = 0$ si ha $0(g \dot{+} g') = \dot{0} = \dot{0} \dot{+} \dot{0}$ (assiomi di gruppo) $= 0g \dot{+} 0g'$. Per $n = 1$ si ha $1(g \dot{+} g') = g \dot{+} g' = 1g \dot{+} 1g'$.

(v4) $(m +_2 n)g = mg \dot{+} ng$: distinguiamo tra due casi. Per $n = 0$ abbiamo $(m +_2 0)g = mg = mg \dot{+} \dot{0}$ (assiomi di gruppo) $= mg \dot{+} 0g$. Per $n = 1$ dobbiamo considerare due sottocasi. Per $m = 0$ abbiamo $(0 +_2 1)g = 1g = \dot{0} \dot{+} 1g$ (assiomi di gruppo) $= 0g \dot{+} 1g$. Per $m = 1$ abbiamo $(1 +_2 1)g = 0g = \dot{0} = g \dot{+} g$ (assioma (ord2)) $= 1g \dot{+} 1g$.

Pertanto i modelli di $T$ sono esattamente i gruppi abeliani che sono spazi vettoriali su $\mathbb{Z}_2$, e pertanto $T$ avrà modelli di qualsiasi cardinalità $k \geq \aleph_0$.

Vediamo ora che $T$ è $k$-categorica per $k \geq \aleph_0$. Siano infatti dati due gruppi abeliani $\mathcal{G}, \mathcal{G}'$ che siano spazi vettorial su $\mathbb{Z}_2$ e di cardinalità

$k \geq \aleph_0$. Poiché $\mathbb{Z}_2$ è finito, per poter ottenere (esattamente) $k$ combinazioni lineari distinte in forma normale e a coefficienti non nulli in $\mathbb{Z}_2$ occorre che $\mathscr{G}$ e $\mathscr{G}'$ abbiano basi di cardinalità $k$, cioè abbiano dimensione $k$. Per il Teorema della Dimensione 3.20, $\mathscr{G}$ e $\mathscr{G}'$ saranno isomorfi in quanto spazi vettoriali. Analogamente alla dimostrazione del Teorema 6.3 saranno isomorfi anche in quanto gruppi sul linguaggio $\mathscr{L}$. $\square$

# Capitolo 7

# Eliminazione dei quantificatori

## *7.1 Teorie con eliminazione dei quantificatori*

**Definizione 7.1.** *Una $\mathscr{L}$-teoria $T$ ha l'*eliminazione dei quantificatori *se per ogni $\mathscr{L}$-formula $\varphi(x_1, ..., x_n)$ esiste una $\mathscr{L}$-formula $\psi(x_1, ..., x_n)$ con le stesse variabili libere e priva di quantificatori tale che*

$$T \vdash \varphi(x_1, ..., x_n) \leftrightarrow \psi(x_1, ..., x_n).$$

**Procedimento standard per l'eliminazione dei quantificatori in una $\mathscr{L}$-formula $\varphi$:**

- riscriviamo $\varphi$ in forma normale rettificata $\varphi'$
- riscriviamo $\varphi'$ in forma normale prenessa

  $$\varphi'' := (Q_1 y_1)...(Q_m y_m)\psi$$

  con $Q_i \in \{\forall, \exists\}$ per $1 \leq i \leq m$ e $\psi$ $\mathscr{L}$-formula senza quantificatori
- cominciamo ad eliminare i quantificatori cominciando dal quantificatore più interno $(Q_m y_m)$. Dobbiamo cioè trovare una formula $\psi'$ senza quantificatori tale che

  $$T \vdash (Q_m y_m)\psi \leftrightarrow \psi'.$$

Senza perdita di generalità si può supporre che $(Q_m y_m) \equiv (\exists y_m)$. Infatti siccome $\vdash (\forall y_m)\psi \leftrightarrow \neg(\exists y_m)\neg\psi$, se $(Q_m y_m) \equiv (\forall y_m)$ è sufficiente eliminare $(\exists y_m)$ da $(\exists y_m)\neg\psi$ ottenendo così $\xi$ senza quantificatori tale che $T \vdash (\exists y_m)\neg\psi \leftrightarrow \xi$, e cioè $T \vdash (\forall y_m)\psi \leftrightarrow \neg\xi$. A questo punto basta porre $\psi' :\equiv \neg\xi$. Sia quindi $(Q_m y_m) \equiv (\exists y_m)$

- scriviamo $\psi$ in forma nomale disgiuntiva $\bigvee_{i=1}^{k}\left(\bigwedge_{j=1}^{l(i)} L_{ij}\right)$ con ciascun $L_{ij}$ letterale (compresi qui possibilmente $\top$ e $\bot$, per $\top \equiv_{def} \neg\bot$) utilizzando le equivalenze $\vdash \neg\neg\chi \leftrightarrow \chi$, $\vdash (\chi \to \xi) \leftrightarrow \neg\chi \vee \xi$, $\vdash \neg(\chi\Box\xi) \leftrightarrow \neg\chi\Diamond\neg\xi$, $\vdash \chi\Box(\xi\Diamond\zeta) \leftrightarrow (\chi\Box\xi)\Diamond(\chi\Box\zeta)$ per $\Box, \Diamond \in \{\wedge, \vee\}$ e $\Box \not\equiv \Diamond$

- se $\mathscr{L}$ non contiene costanti individuali, rimpiazziamo ogni letterale $L_{ij}(y_m)$ tale che $T \vdash (\forall y_m)L_{ij}(y_m)$ con $\top$; infatti $T \vdash L_{ij}(y_m) \leftrightarrow \top$. Ogni letterale $L_{ij}(y_m)$ tale che $T \vdash (\forall y_m)\neg L_{ij}(y_m)$ viene invece rimpiazzato con $\bot$: infatti $T \vdash L_{ij}(y_m) \leftrightarrow \bot$. L'idea alla base di queste sostituzioni è che se $L_{ij}$ contiene al massimo $y_m$, tale letterale non contiene variabili libere all'interno di $(\exists y_m)\left(\bigvee_{i=1}^{k}\left(\bigwedge_{j=1}^{l(i)} L_{ij}\right)\right)$ essendo vincolato da $(\exists y_m)$, quindi se alla fine questo quantificatore viene eliminato conviene sostituire $L_{ij}(y_m)$ con una formula senza variabili per essere sicuri che la condizione sulle variabili prevista dalla Definizione 7.1 venga rispettata. Otteniamo così i letterali $M_{i_j}$ che coincidono con i precedenti $L_{i_j}$ salvo che per alcune loro eventuali sostituzioni con $\top$ e $\bot$ come indicato

- osserviamo che

$$\vdash (\exists y_m)\psi \leftrightarrow (\exists y_m)\left(\bigvee_{i=1}^{k}\left(\bigwedge_{j=1}^{l(i)} M_{ij}\right)\right) \leftrightarrow \bigvee_{i=1}^{k}(\exists y_m)\left(\bigwedge_{j=1}^{l(i)} M_{ij}\right)$$

(si ricordi che i quantificatori esistenziali commutano infatti perfettamente con $\vee$ ma non con $\wedge$). Siano inoltre $M_{ij_1}, .., M_{ij_{t(i)}}$ i letterali in $\{M_{i1} ..., M_{il(i)}\}$ che contengono $y_m$, e $M'_{ij_1}, .., M'_{ij_{p(i)}}$ i letterali in $\{M_{i1} ..., M_{il(i)}\}$ che non contengono $y_m$, per $1 \leq i \leq$

$k$. Allora osserviamo che

$$\vdash (\exists y_m)\left(\bigwedge_{j=1}^{l(i)} M_{ij}\right) \leftrightarrow (\exists y_m)\left(\bigwedge_{s=1}^{t(i)} M_{ij_s}\right) \wedge \bigwedge_{s=1}^{p(i)} M'_{ij_s}.$$

Ci riduciamo allora ad eliminare $(\exists y_m)$ nelle formule

$$(\exists y_m)\left(\bigwedge_{s=1}^{t(i)} M_{ij_s}\right).$$

- ripetiamo il procedimento muovendoci verso sinistra fino a raggiungere $(Q_1 y_1)$.

**Teorema 7.2.** *La teoria degli Ordini Lineari Densi Stretti senza Massimo né Minimo ha l'eliminazione dei quantificatori.*

*Dimostrazione.* Osservando che per l'assioma di linearità $T \vdash x \neq z \leftrightarrow (x < z \vee z < x)$, $T \vdash x \not< z \leftrightarrow (x = z \vee z < x)$ ci riduciamo ad eliminare $(\exists y)$ in formule del tipo

$$(\exists y)\left(\underbrace{\bigwedge_{i=1}^{n} y = v_i}_{(a)} \wedge \underbrace{\bigwedge_{j=1}^{k} y < u_j}_{(b)} \wedge \underbrace{\bigwedge_{h=1}^{l} w_h < y}_{(c)}\right)$$

per $v_i, u_j, w_h$ variabili diverse da $y$ (la nostra teoria non contiene né costanti individuali né simboli funtoriali), e dove ovviamente qualcuna delle parti (a), (b), (c) può mancare. Abbiamo quindi i seguenti casi:

- sono presenti (a), (b), (c): la formula si riduce a:

$$\left(\bigwedge_{i=1}^{n} v_1 = v_i\right) \wedge \left(\bigwedge_{j=1}^{k} v_1 < u_j\right) \wedge \left(\bigwedge_{h=1}^{l} w_h < v_1\right).$$

- (a) è presente ma manca (b) oppure (c) (possibilmente entrambi): segue in modo banale dal caso precedente

- solo (b) è presente: la formula è vera in ogni modello di $T$ in quanto tali modelli non hanno minimo. Possiamo quindi prendere una formula valida in tutti i modelli di $T$ nelle stesse variabili libere, anzi possiamo prendere una formula logicamente valida nelle stesse variabili libere:

$$\bigwedge_{j=1}^{k} u_j = u_j$$

- solo (c) è presente: la formula è vera in ogni modello di $T$ in quanto tali modelli non hanno massimo. Possiamo quindi prendere una formula valida in tutti i modelli di $T$ nelle stesse variabili libere, anzi possiamo prendere una formula logicamente valida nelle stesse variabili libere:

$$\bigwedge_{h=1}^{l} w_h = w_h$$

- (a) manca ma sia (b) che (c) sono presenti: mostriamo che la formula si riduce a

$$\bigwedge_{j=1}^{k} \bigwedge_{h=1}^{l} w_h < u_j.$$

  Da un lato, sia infatti $\mathscr{M} \models (\exists y)\Big((\bigwedge_{j=1}^{k} y < u_j) \wedge (\bigwedge_{h=1}^{l} w_h < y)\Big)[(m_j)_{1\leq j\leq k}, (m'_h)_{1\leq h\leq l}]$ per $\mathscr{M}$ modello di $T$ con domino $M$. Allora esiste un $m \in M$ tale che $m'_h < m$ e $m < m_j$ per $1 \leq j \leq k$, $1 \leq h \leq l$. Da questo segue per transitività $m'_h < m_j$ per $1 \leq j \leq k$, $1 \leq h \leq l$. Per cui

$$\mathscr{M} \models \bigwedge_{j=1}^{k} \bigwedge_{h=1}^{l} w_h < u_j[(m_j)_{1\leq j\leq k}, (m'_h)_{1\leq h\leq l}].$$

  Dall'altro lato sia

$$\mathscr{M} \models \bigwedge_{j=1}^{k} \bigwedge_{h=1}^{l} w_h < u_j[(m'_h)_{1\leq h\leq l}, (m_j)_{1\leq j\leq k}]$$

per $\mathscr{M}$ modello di $T$ con domino $M$. Sia $m_J$ il minimo degli $m_j$ per $1 \leq j \leq k$ ed $m'_H$ il massimo degli $m'_h$ per $1 \leq h \leq l$ (tali elementi esistono perché l'ordine è lineare). Abbiamo che $m'_H < m_J$. Giacché $\mathscr{M}$ è un ordine denso, allora esiste $m \in M$ con $m'_H < m < m_J$, da cui $m'_h < m < m_j$ per $1 \leq h \leq l$, $1 \leq j \leq k$ per transitività. Perciò $\mathscr{M} \models (\exists y)\Big((\bigwedge_{j=1}^{k} y < u_j) \wedge (\bigwedge_{h=1}^{l} w_h < y)\Big)[(m'_h)_{1 \leq h \leq l}, (m_j)_{1 \leq j \leq k}]$.

□

**Esempio 7.3.** *Consideriamo la formula*

$$\varphi :\equiv (\forall x)(\exists y)(\exists z)(x < y \wedge z < x).$$

*e applichiamo ad essa l'algoritmo di eliminazione dei quantificatori*[1]. *Abbiamo che* $\vdash (\exists z)(x < y \wedge z < x) \leftrightarrow (x < y \wedge (\exists z)\, z < x)$. *La formula* $(\exists z) z < x$ *si riduce a* $x = x$. *Come primo passaggio otteniamo quindi*

$$(\forall x)(\exists y)(x < y \wedge x = x).$$

*Abbiamo che* $\vdash (\exists y)(x < y \wedge x = x) \leftrightarrow (\exists y)\, x < y \wedge x = x$. *La formula* $(\exists y) x < y$ *si ridurrà a* $x = x$. *Come secondo passaggio otteniamo quindi* $(\forall x)(x = x \wedge x = x)$, *o più semplicemente*

$$(\forall x)\, x = x.$$

*Sappiamo che* $\vdash (\forall x) x = x \leftrightarrow \neg(\exists x)\neg\, x = x$. *Consideriamo pertanto* $(\exists x)\neg\, x = x$. *Abbiamo che* $\vdash (\forall x)\neg x = x \leftrightarrow \bot$, *per cui riduciamo* $\neg x = x$ *a* $\bot$ *e cancelliamo* $(\exists x)$. *In conclusione* $\varphi$ *viene ridotta a* $\neg\bot$, *cioè a* $\top$.

Consideriamo ora la teoria dei Campi Algebricamente Chiusi (Algebraically Closed Fields) nel linguaggio dei campi $\mathscr{L} := \{\dot{0}, \dot{1}, \dot{-}, \dot{+}, *\}$. Questa è data dagli assiomi dei campi più l'assioma (schema di assioma) seguente

[1]Essendo tale formula un assioma della teoria, il risultato finale è ovvio, ma lo scopo dell'esempio è quello di applicare l'algoritmo di eliminazione dei quantificatori passo dopo passo.

(acf) $(\forall x_0)\ldots(\forall x_n)(\exists y)(x_n \neq \dot{0} \rightarrow x_n y^n \dot{+} x_{n-1} y^{n-1} \dot{+} \ldots \dot{+} x_1 y \dot{+} x_0 = \dot{0})$

per $n \geq 1$.

**Teorema 7.4.** *La teoria dei Campi Algebricamente Chiusi (ACF) ha l'eliminazione dei quantificatori.*

Una dimostrazione completa di questo teorema esula dagli obiettivi di questo volume (si veda in merito Marcja & Toffalori [2003]). Vediamo tuttavia come si potrebbero eliminare i quantificatori in un caso rilevante. Innanzitutto osserviamo che è sempre sufficiente considerare formule del tipo

$$(\exists y)\left(\bigwedge_{i=1}^{k} t_i = \dot{0} \wedge \bigwedge_{j=1}^{l} t'_j \neq \dot{0}\right)$$

senza quantificatori all'interno del campo di azione di $(\exists y)$, dove ciascun $t_i, t'_j$, è un (termine-somma esprimente un) polinomio in $y$. Mediante usuali calcoli algebrici possiamo supporre che ciascun $t_i$, rispettivamente $t'_j$, sia in forma normale polinomiale

$$a_n y^n \dot{+} a_{n-1} y^{n-1} \dot{+} \ldots \dot{+} a_1 y \dot{+} a_0$$

per $a_n, \ldots, a_0$ $\mathscr{L}$-termini che non contengono $y$ (per $i \geq 0$ si ponga ancora $y^i$ come abbreviazione notazionale per $\underbrace{y * \ldots * y}_{i\text{-volte}}$ ed inoltre si ponga $ss'$ come abbreviazione notazionale per $s * s'$ per $s, s'$ $\mathscr{L}$-termini)[2]. Possiamo anche assumere sempre che sia $n > 0$: in caso contrario di fatto la $y$ non occorre in $t_i$, rispettivamente $t'_j$, e sappiamo già che porteremo $t_i = \dot{0}$, rispettivamente $t'_j \neq \dot{0}$, fuori dal campo di azione del quantificatore.

Analizziamo solo il caso in cui dobbiamo eliminare un solo quantificatore esistenziale $(\exists y)$ da un sistema di equazioni contenente

[2]Tali calcoli algebrici sono giustificati dall'interpretazione della $y$ sugli elementi del dominio $K$ di un generico campo $\mathscr{K}$, ovvero dall'interpretazione dei polinomi come funzioni polinomiali. Ad esempio, per le leggi dei campi per $k \in K$ si ha che $a * k^i \dot{+} b * k^i = \dot{0}$ sse $(a \dot{+} b) * k^i = \dot{0}$ (distributività), pertanto $(\exists y) a y^i \dot{+} b y^i = \dot{0}$ è vera in un campo sse lo è $(\exists y)(a \dot{+} b) y^i = \dot{0}$. Analogamente $(\exists y) a_1 y^m \dot{+} \ldots \dot{+} a_{i+1} y^{i+1} \dot{+} a_{i-1} y^{i-1} \dot{+} \ldots = \dot{0}$ è vera sse lo è $(\exists y) a_1 y^m \dot{+} \ldots \dot{+} a_{i+1} y^{i+1} \dot{+} \dot{0}\, y^i \dot{+} a_{i-1} y^{i-1} \dot{+} \ldots = \dot{0}$, e così via.

la sola variabile $y$ (pertanto i termini esprimenti i coefficienti delle potenze $i$-esime di $y$, per $i \geq 0$, sono tutti $\mathscr{L}$-termini chiusi).

Distinguiamo ora tra due sottocasi:

Caso $k = 1$) dobbiamo eliminare il quantificatore da $(\exists y)\, t_1 = \dot{0}$. Vediamo $t_1$ come $t_1 := a_n y^n \dot{+} r_1$, per $n > 0$, dove $r_1$ è (un termine-somma esprimente) un polinomio con $n - 1$ come massima potenza della $y$. Si danno due casi: o $a_n \neq \dot{0}$, nel qual caso per l'assioma (acf) il polinomio $t_1$ ha (almeno) una radice, oppure $a_n = \dot{0}$, nel qual caso ci ridurremo a chiederci se $r_1$ ha una radice. Pertanto $(\exists y)\ t_1 = \dot{0}$ equivale alla formula

$$a_n \neq \dot{0} \vee \Big(a_n = \dot{0} \wedge (\exists y)\, r_1 = \dot{0}\Big).$$

Possiamo ora ripetere il procedimento per eliminare il quantificatore esistenziale da $(\exists y)\, r_1 = \dot{0}$, ricordando che $r_1$ ha massima potenza della $y$ inferiore a $t_1$. La massima potenza di $y$ decrescerà quindi ad ogni passo di iterazione del procedimento fino a raggiungere lo zero (cioè, la $y$ scomparirà). A questo punto, portare la corrispondente equazione fuori dal quantificatore $(\exists y)$ significa ovviamente semplicemente cancellare $(\exists y)$ in quanto vacuo. Poiché $\vdash \neg\chi \vee (\chi \wedge \xi) \leftrightarrow \neg\chi \vee \xi$, l'enunciato $(\exists y)\ t_1 = \dot{0}$ si riduce alla fine a

$$a_n \neq \dot{0} \vee \ldots \vee a_1 \neq \dot{0} \vee a_0 = \dot{0}.$$

Si noti che i valori di verità di questi disgiunti cambieranno in generale a seconda delle caratteristiche dei campi.

Caso $k > 1$) consideriamo i primi due polinomi nella congiunzione, e siano questi, analogamente al caso precedente, $t_1 :\equiv a_n y^n \dot{+} r_1$ e $t_2 :\equiv b_m y^m \dot{+} r_2$. Senza perdita di generalità possiamo assumere che $n \geq m > 0$. Distinguiamo due casi a seconda che $b_m$ si annulli oppure no (il che generalmente dipenderà dalle caratteristiche dei campi). Nel primo caso la domanda sull'esistenza di una radice in comune si restringe a $t_1$ ed $r_2$. Nel secondo caso invece consideriamo il primo passaggio della divisione euclidea di $t_1$ per $t_2$ (essendo infatti $b_m \neq \dot{0}$ tale passaggio può essere eseguito), e la domanda sull'esistenza di una radice in comune viene spostata sul primo resto della divisione e su $t_2$. In sintesi, $t_1 = \dot{0} \wedge t_2 = \dot{0}$ viene rimpiazzata con

$$(b_m = \dot{0} \wedge t_1 = \dot{0} \wedge r_2 = \dot{0}) \vee (b_m \neq \dot{0} \wedge t_1' = \dot{0} \wedge t_2 = \dot{0})$$

per $t_1'$ il polinomio corrispondente a $b_m t_1 \dot{-} a_n y^{n-m} t_2$ (cioè $t_1'$ corrisponde al resto del primo passaggio della divisione di $t_1$ per $t_2$ moltiplicato per il denominatore del coefficiente del primo quoziente provvisorio della divisione, cioè il primo coefficiente di $t_2$, ovvero $b_m$; si noti in merito quanto da noi osservato nella Sezione 3.1.4 sulla divisione dei polinomi). Poiché $b_m$ è un termine non contenente $y$, otteniamo

$$\left(b_m = \dot{0} \wedge (\exists y)\left(t_1 = \dot{0} \wedge r_2 = \dot{0} \wedge \bigwedge_{i=3}^{k} t_i = \dot{0}\right)\right) \vee$$

$$\vee \left(b_m \neq \dot{0} \wedge (\exists y)\left(t_1' = \dot{0} \wedge t_2 = \dot{0} \wedge \bigwedge_{i=3}^{k} t_i = \dot{0}\right)\right).$$

Ci siamo ridotti in entrambi i disgiunti ad eliminare $(\exists y)$ da congiunzioni di equazioni in cui un termine ha diminuito la propria massima potenza in $y$ (infatti $r_2$ ha potenza massima inferiore rispetto a $t_2$ e $t_1'$ ha potenza massima inferiore rispetto a $t_1$). Ripetiamo ora il procedimento in entrambi i disgiunti rispetto a $t_1, r_2$ e $t_1', t_2$ rispettivamente, e così via, facendo calare man mano le potenze della $y$.

Ogni volta che si raggiunge la 0-esima potenza, ovvero ci si riduce ad una semplice costante, la $y$ di fatto scompare, e il quantificatore $(\exists y)$ diventa inutile; possiamo perciò cancellarlo[3].

Si procede in questo modo fino a ridursi ad un solo polinomio con quantificatore esistenziale (m.c.d.) che verrà trattato come nel caso $k = 1$.

[3] Si osservi che qualora $t_2$ corrisponda ad un polinomio nullo, l'algoritmo riduce la domanda sull'esistenza di radici in comune tra $t_1$ e $t_2$ semplicemente a quella sull'esistenza delle radici del solo $t_1$, il che è coerente col fatto che ogni oggetto del campo sarà una radice del polinomio nullo.

**Esercizio 7.5.** *Mediante il metodo per l'eliminazione dei quantificatori stabilire se l'equazione* $5y + 10 = 0$ *ha soluzioni in ogni campo algebricamente chiuso*[4].

SOLUZIONE: Abbiamo $(\exists y)\, 5y + 10 = 0 \rightsquigarrow 5 \neq 0 \vee 10 = 0$.

Quindi sì, perché se il campo ha caratteristica diversa da 5, allora $5 \neq 0$. Altrimenti, $5 = 10 = 0 \mod 5$.

**Esercizio 7.6.** *Si consideri la* teoria dei campi algebricamente chiusi di caratteristica $p$ *ottenuta aggiungendo agli assiomi dei campi algebricamente chiusi l'assioma*

$$(cp)\quad \underbrace{\dot{1} \dot{+} \ldots \dot{+} \dot{1}}_{p\text{-volte}} = \dot{0}$$

*e la* teoria dei campi algebricamente chiusi di caratteristica 0 *ottenuta aggiungendo agli assiomi dei campi algebricamente chiusi l'assioma (schema di assioma)*

$$(c0)\quad \underbrace{\dot{1} \dot{+} \ldots \dot{+} \dot{1}}_{n\text{-volte}} \neq \dot{0}$$

*per* $n \geq 1$. *Allora:*

1. *con il metodo esposto per l'eliminazione dei quantificatori stabilisci se l'equazione* $2y + 5 = 0$ *ha soluzioni*

   (a) *nei campi algebricamente chiusi di caratteristica* 0

   (b) *nei campi algebricamente chiusi di caratteristica* 2

2. *applica il metodo per l'eliminazione dei quantificatori ai seguenti sistemi di equazioni per stabilire se abbiano soluzione nei campi algebricamente chiusi di caratteristica* 0 *e confronta tali risultati con quelli dell'Esercizio 3.9:*

   (a) $2y = 0, 4y = 0$

   (b) $4y^2 + 2y = 0, 5y + 10 = 0$

   (c) $27y^3 - 75y = 0, 3y^2 + 8y + 5 = 0$

[4] In questo esercizio e nel prossimo denoteremo i coefficienti numerici mediante numerali usuali anziché termini chiusi nel linguaggio $\mathscr{L} = \{\dot{0}, \dot{1}, \dot{-}, \dot{+}, *\}$.

*SOLUZIONI:*

1. (a) $(\exists y)\, 2y+5=0 \rightsquigarrow 2\neq 0$. *Sì.*
   (b) $(\exists y)\, 2y+5=0 \rightsquigarrow 2\neq 0 \vee 5=0$. *No.*

2. (a) $(\exists y)(2y=0\wedge 4y=0) \rightsquigarrow (\exists y)(4\neq 0\wedge 8y-8y=0\wedge 4y=0) \rightsquigarrow$ $(\exists y)(4\neq 0 \wedge 0=0\wedge 4y=0) \rightsquigarrow 4\neq 0\wedge 0=0\wedge (\exists \mathbf{y})\, \mathbf{4y}=\mathbf{0} \rightsquigarrow$ $\rightsquigarrow 4\neq 0 \wedge 0=0 \wedge \mathbf{4}\neq\mathbf{0}$. *L'enunciato* $(\exists y)(2y=0\wedge 4y=0)$ *corrisponde a chiedersi se il m.c.d.* $4y$ *(di* $2y$ *e* $4y$*) abbia delle soluzioni. In un campo algebricamente chiuso di caratteristica* $0$ *questo è immediatamente verificato dal fatto che* $4\neq 0$*. Quindi sì.*
   (b) $(\exists y)(4y^2+2y=0 \wedge 5y+10=0) \rightsquigarrow$
   $\rightsquigarrow 5\neq 0\wedge(\exists y)[20y^2+10y-4y(5y+10)=0\wedge 5y+10=0] \rightsquigarrow$
   $\rightsquigarrow (\exists y)[-30y=0\wedge 5y+10=0] \rightsquigarrow$
   $\rightsquigarrow 5\neq 0\wedge (\exists y)[-150y+30(5y+10)=0\wedge 5y+10=0] \rightsquigarrow$
   $\rightsquigarrow \mathbf{300}=\mathbf{0}\wedge(\exists y)5y+10=0$. *Durante i passaggi otteniamo dei polinomi che corrispondono ai resti rispetto alle divisioni per* $5y+10$ *trovati nell'Esercizio 3.9 moltiplicati per potenze del coefficiente del primo addendo di tale polinomio (*$5$*). La costante* $300$ *è un multiplo (rispetto al quadrato di* $5$*) del m.c.d. 12 trovato nell'Esercizio 3.9, che è una costante diversa da* $0$*. Quindi il sistema non può avere delle soluzioni, e d'altronde non è vero che* $300=0$.
   (c) $(\exists y)(27y^3-75y=0\wedge \underbrace{3y^2+8y+5}_{B(y)}=0) \rightsquigarrow$
   $\rightsquigarrow 3\neq 0\wedge(\exists y)[81y^3-225y-27y(3y^2+8y+5)=0\wedge B(y)=0] \rightsquigarrow$
   $\rightsquigarrow (\exists y)(-216y^2-360y=0\wedge B(y)=0) \rightsquigarrow$
   $\rightsquigarrow (\exists y)[-648y^2-1080+216(3y^2+8y+5)=0\wedge B(y)] \rightsquigarrow$
   $\rightsquigarrow (\exists y)(\underbrace{648y+1080}_{B'(y)}=0\wedge 3y^2+8y+5=0) \rightsquigarrow$
   $\rightsquigarrow 648\neq 0\wedge(\exists y)[1944y^2+5184y+3240-3y(648y+1080)\wedge B'(y)] \rightsquigarrow$
   $\rightsquigarrow (\exists y)(1944y+3240=0\wedge B'(y)) \rightsquigarrow$
   $\rightsquigarrow (\exists y)[1259712y+2099520-1944(648y+1080)\wedge B'(y)=0] \rightsquigarrow$

$\rightsquigarrow 0 = 0 \wedge (\exists \mathbf{y})\mathbf{648y} + \mathbf{1080} = \mathbf{0} \rightsquigarrow \mathbf{648} \neq \mathbf{0}$. *Durante i passaggi otteniamo dei polinomi che corrispondono ai resti rispetto alle divisioni per* $B(y) := 3y^2 + 8y + 5$ *prima, e per* $B'(y) := 648y + 1080$ *poi, moltiplicati per occorrenze dei coefficienti dei primi addendi di tali polinomi (ovvero 3 e 648 rispettivamente). Il polinomio* $648y + 1080$ *è un multiplo (rispetto al quadrato del primo coefficiente di* $B(x)$*) del m.c.d.* $72y + 120$ *trovato nell'Esercizio 3.9, con esso condivide pertanto le stesse radici ed è a sua volta un m.c.d. (di* $27y^3 - 75y$ *e* $3y^2 + 8y + 5$*). Il fatto che tale polinomio abbia soluzioni è verificato dal fatto che* $648 \neq 0$. *Quindi sì, il sistema di equazioni ha soluzioni.*

## *7.2 Insiemi definibili*

**Definizione 7.7** (Insiemi definibili)**.** *Dato un linguaggio* $\mathscr{L}$ *ed una* $\mathscr{L}$*-struttura* $\mathscr{A}$ *con dominio A, un insieme* $D \subseteq A$ *è* definibile *in* $\mathscr{A}$ *se esiste una* $\mathscr{L}$*-formula* $\varphi(x, x_1, ..., x_n)$ *ed esistono* $a_1, ..., a_n \in A$ *tali che*

$$D := \left\{ a \in A \mid \mathscr{A} \models \varphi[a, a_1, ..., a_n] \right\}.$$

**Definizione 7.8** (combinazioni booleane)**.** *Sia dato un insieme* $X$ *e siano* $A, A_1, ..., A_n \subseteq X$. *Allora* $A$ *è* combinazione booleana *di* $A_1, ..., A_n$ *se è ottenibile applicando ad* $A_1, ..., A_n$ *operazioni in* $\{\cap, \cup, X \setminus \cdot\}$.

**Definizione 7.9** (insiemi algebrici)**.** *Sia dato un campo algebricamente chiuso* $\mathscr{K}$ *con dominio* $K$. *Allora* $V \subseteq K$ *è un* insieme algebrico *se costituisce l'insieme di soluzioni di un sistema di equazioni algebriche su* $K$ *in una stessa variabile* $x$, *dove un'equazione algebrica è un'equazione del tipo* $A(x) = \dot{0}$ *per* $A(x)$ *polinomio in* $x$.

**Corollario 7.10.** *Gli insiemi definibili in* $\mathbb{C}$ *sono esattamente le combinazioni booleane di insiemi algebrici. In particolare gli insiemi definibili in* $\mathbb{C}$ *sono finiti o cofiniti.*

*Dimostrazione.* La $\mathscr{L}$-teoria

$$Th(\mathbb{C}) = \left\{ \varphi \mid \varphi \ \mathscr{L}\text{-formula chiusa e } \mathbb{C} \models \varphi \right\}$$

per $\mathscr{L} := \{\dot{0}, \dot{1}, \dot{-}, \dot{+}, *\}$ include la teoria dei Campi Algebricamente Chiusi (in quanto $\mathbb{C}$ è un campo algebricamente chiuso[5]). Possiamo allora vedere che $Th(\mathbb{C})$ ha l'eliminazione dei quantificatori. Sia infatti $\varphi(x_1,...,x_n)$ una $\mathscr{L}$-formula chiusa. Poiché $ACF$ ha l'eliminazione dei quantificatori per Teorema 7.4, esiste una $\mathscr{L}$-formula $\psi(x_1,...,x_n)$ senza quantificatori tale che

$$ACF \vdash \varphi(x_1,...,x_n) \leftrightarrow \psi(x_1,...,x_n).$$

Ma $ACF \subseteq Th(\mathbb{C})$, da cui $Th(\mathbb{C}) \vdash \varphi(x_1,...,x_n) \leftrightarrow \psi(x_1,...,x_n)$.

Sia ora $D \subseteq \mathbb{C}$ definito dalla formula $\varphi(x, x_1,...,x_n)$ mediante gli oggetti $c_1,..,c_n \in \mathbb{C}$. Poiché, come abbiamo appena visto, $Th(\mathbb{C})$ ha l'eliminazione dei quantificatori, esiste una formula $\psi(x, x_1,...,x_n)$ senza quantificatori, tale che $Th(\mathbb{C}) \vdash \varphi(x, x_1,...,x_n) \leftrightarrow \psi(x, x_1,...,x_n)$. Pertanto $D = \left\{c \in \mathbb{C} \mid \mathbb{C} \models \varphi[c, c_1,...,c_n]\right\} = \left\{c \in \mathbb{C} \mid \mathbb{C} \models \psi[c, c_1,...,c_n]\right\}$, ovvero $D$ è definito anche da $\psi(x, x_1,...,x_n)$ mediante $c_1,...,c_n$. Per l'algoritmo (generale) di eliminazione dei quantificatori, possiamo supporre che $\psi$ sia una forma normale disgiuntiva le cui formule atomiche rappresentano equazioni algebriche. Queste formule sono pertanto del tipo $t = 0$, dove però potrebbe non essere univocamente determinata la variabile secondo la quale il polinomio (espresso da) $t$ è inteso. Infatti $t$ potrebbe contenere più variabili in $\{x_1,...,x_n\}$, Tuttavia, nel momento in cui $x_1,...,x_n$ vengono interpretate su $c_1,...,c_n$, allora otteniamo di fatto un polinomio nella sola variabile $x$. In conclusione, $D$ sarà una combinazione booleana di insiemi algebrici definiti da equazioni nella variabile $x$ mediante l'interpretazione dei connettivi $\neg, \wedge, \vee$ sulle operazioni insiemistiche $\mathbb{C} \setminus ., \cap, \cup$ rispettivamente.

Viceversa, se $D$ è una combinazione booleana di insiemi algebrici, allora è facile vedere che tale insieme è definibile in $\mathbb{C}$ (considerando le equazioni che definiscono tali insiemi algebrici come espresse in una medesima variabile, e mediante la corrispondenza tra le operazioni insiemistiche di base e i connettivi proposizionali booleani appena espressa).

Poiché un'equazione algebrica con parametri in $\mathbb{C}$ definisce un insieme finito (se del tipo $t = 0$ con $t$ polinomio non nullo), o cofinito (se del tipo $t = 0$ con $t$ polinomio nullo, nel qual caso definisce tutto

[5]Di fatto, prendiamo come $Th(\mathbb{C})$ la teoria dei Campi Algebricamente Chiusi di Caratteristica 0, vedasi l'Esercizio 7.6.

$\mathbb{C}$), e poiché le combinazioni booleane di insiemi finiti e cofiniti sono ancora insiemi finiti e cofiniti, $D$ sarà un insieme finito o cofinito. □

**Corollario 7.11.** *In* $\mathbb{C}$ *gli insiemi* $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ *non sono definibili.*

*Dimostrazione.* Segue dal Corollario 7.10, in quanto gli insiemi $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ non sono né finiti né cofiniti. □

Le teorie che hanno l'eliminazione dei quantificatori sono matematicamente molto comode perché di fatto l'aumento della complessità logica nella definizione di un insieme è solamente apparente!

# Capitolo 8

# Teorie Decidibili

**Definizione 8.1** (Teorie decidibili)**.** *Una $\mathscr{L}$-teoria $T$ è* decidibile *quando esiste una procedura meccanica (algoritmo) che, data una qualsiasi $\mathscr{L}$-formula (chiusa)*[1]*, dia risposta positiva se $\varphi$ è teorema di $T$ ($T \vdash \varphi$) e risposta negativa se $\varphi$ non lo è ($T \nvdash \varphi$).*

**Definizione 8.2** (Teorie complete)**.** *Una $\mathscr{L}$-teoria $T$ è* completa *quando per qualsiasi $\mathscr{L}$-formula chiusa $\varphi$, o $\varphi$ stessa è teorema di $T$ ($T \vdash \varphi$) oppure lo è la sua negazione ($T \vdash \neg\varphi$).*

Anche il seguente teorema, come quello di Compattezza e quelli di Skolem-Löwenheim, può essere formulato per cardinalità infinite arbitrariamente grandi, ma noi ci soffermeremo solo sul caso dell'infinito numerabile:

**Teorema 8.3** (Test di Vaught (caso numerabile))**.** *Per $\mathscr{L}$ numerabile, sia $T$ una $\mathscr{L}$-teoria soddisfacibile che ha solo modelli infiniti e che è $\aleph_0$-categorica. Allora $T$ è completa.*

*Dimostrazione.* Sia per assurdo $\varphi$ una $\mathscr{L}$-formula chiusa tale che $T \nvdash \varphi$ e $T \nvdash \neg\varphi$. Allora per Teorema di Completezza sia $T \cup \{\neg\varphi\}$ che $T \cup \{\varphi\}$ hanno un modello. Poiché $T$ possiede solamente modelli infiniti, per il Teorema di Skolem-Löwenheim verso il basso 4.6, $T \cup \{\neg\varphi\}$ e $T \cup \{\varphi\}$ possiedono modelli infiniti numerabili $\mathscr{M}$ ed $\mathscr{M}'$

[1] Possiamo sempre assumere che la formula sia chiusa mediante chiusura universale delle variabili libere.

rispettivamente. Poichè $T$ è $\aleph_0$-categorica, $\mathscr{M} \approx \mathscr{M}'$, da cui $\mathscr{M} \equiv \mathscr{M}'$ per Osservazione 5.12, e quindi dovrebbero soddisfare entrambi sia $\varphi$ che $\neg\varphi$, il che è assurdo. □

**Corollario 8.4.** *La teoria degli Ordini Lineari Stretti Densi senza Massimo né Minimo è completa.*

*Dimostrazione.* Segue dal fatto che ha solo modelli infiniti e dalla sua $\aleph_0$-categoricità (Teorema 6.2) applicati al Test di Vaught (Teorema 8.3). □

**Definizione 8.5** (Teorie ricorsivamente assiomatizzabili)**.** *Una $\mathscr{L}$-teoria $T$ è* ricorsivamente assiomatizzabile *se esiste una procedura meccanica (algoritmo) che produca un sistema di assiomi per la teoria.*

Tutti gli esempi di teorie algebriche da noi analizzate sono ovviamente ricorsivamente assiomatizzabili.

**Proposizione 8.6.** *Una $\mathscr{L}$-teoria ricorsivamente assiomatizzabile e completa è decidibile.*

*Dimostrazione.* Se $T$ è inconsistente l'enunciato è banalmente vero. Il caso interessante si dà quindi quando $T$ è consistente, cioè $T \nvdash \perp$. Supponiamo di voler stabilire quando la $\mathscr{L}$-formula chiusa $\varphi$ è un teorema di $T$ e quando non lo è. Giacché $T$ è ricorsivamente assiomatizzabile possiamo generare automaticamente tutti i suoi teoremi. Se tra i teoremi compare $\varphi$ stessa allora ovviamente $T \vdash \varphi$. Se invece vi compare $\neg\varphi$, allora $T \nvdash \varphi$. □

**Corollario 8.7.** *La teoria degli Ordini Lineari Stretti Densi senza Massimo né Minimo è decidibile.*

*Dimostrazione.* Segue dalla sua completezza (Corollario 8.4) e dalla Proposizione 8.6. □

Un esempio di teoria non completa è quella dei Campi Algebricamente Chiusi. Si prenda infatti ad esempio l'enunciato $\dot{1} \dotplus \dot{1} \dotplus \dot{1} = \dot{0}$. Questo enunciato è soddisfatto nei campi algebricamente chiusi di caratteristica 3. Tuttavia per il Teorema di Validità non può essere teorema della teoria, in quanto non è soddisfatto nei campi di caratteristica differente, ad esempio $\mathbb{C}$. Al contempo, neppure $\neg\, \dot{1} \dotplus \dot{1} \dotplus \dot{1} =$

$\dot{0}$ può essere teorema della teoria sempre per il Teorema di Validità, in quanto tale enunciato è falso proprio nei campi di caratteristica 3.

Cionostante si può dimostrare che tale teoria è comunque decidibile mediante la seguente proposizione:

**Proposizione 8.8.** *Sia $T$ una $\mathscr{L}$-teoria per la quale:*

1. *esiste un metodo meccanico per l'eliminazione dei quantificatori in qualsiasi $\mathscr{L}$-formula $\varphi$*[2]
2. *esiste una procedura meccanica che data una formula $\varphi$ chiusa e senza quantificatori dia risposta positiva se essa è un teorema, e risposta negativa se non lo è.*

*Allora $T$ è decidibile.*

La Proposizione 8.8 ci dà innanzitutto una prova alternativa della decidibilità della teoria degli Ordini Lineari Stretti Densi senza Massimo né Minimo:

**Corollario 8.9.** *La teoria degli Ordini Lineari Stretti Densi senza Massimo né Minimo è decidibile.*

*Dimostrazione.* Per l'enunciato del Teorema 7.2 sappiamo che tale teoria $T$ gode della proprietà di eliminazione dei quantificatori, pertanto presa $\varphi(x_1,...,x_n)$ esiste una formula $\psi(x_1,...,x_n)$ senza quantificatori tale che $T \vdash (\forall x_1)...(\forall x_n)(\varphi(x_1,...,x_n) \leftrightarrow \psi(x_1,...,x_n))$. Essendo $T$ anche ricorsivamente assiomatizzabile possiamo trovare in modo automatico una tale formula $\psi$ (la cosa è evidente per "forza bruta", ma nella dimostrazione dello stesso Teorema 7.2 abbiamo anche mostrato una procedura più diretta). Se $\varphi$ è chiusa, allora $\psi$ è ottenuta mediante combinazione di $\perp$ e $\top$ con connettivi. Mediante definizione verofunzionale dei connnettivi possiamo controllare se tale formula è equivalente a $\top$ o $\perp$. Per il Teorema di Completezza, nel primo caso $T \vdash \varphi$, e per il Teorema di Validità, nel secondo $T \nvdash \varphi$. $\square$

Per quanto riguarda la teoria dei Campi Algebricamente Chiusi, dal Teorema 7.4 e dal fatto che tale teoria è ricorsivamente assiomatizzabile la proprietà 1 della Proposizione 8.8 segue immediatamente; resta solo da far vedere che esiste una procedura meccanica per

[2] Si noti che questa proprietà si dà automaticamente per le teorie con eliminazione dei quantificatori e ricorsivamente assiomatizzabili.

stabilire la teorematicità, o equivalentemente la validità rispetto ai modelli di tale teoria, di una formula chiusa senza quantificatori in $\mathscr{L} := \{\dot{0}, \dot{1}, \dot{-}, \dot{+}, *\}$. Ma questa proprietà può essere facilmente dimostrata, ed in un'ultima analisi si riduce a verificare tale proprietà per la teoria dei Campi (non necessariamente algebricamente chiusi). Infatti una $\mathscr{L}$-formula chiusa senza quantificatori ha come formule atomiche equazioni contenenti solo termini chiusi. Senza perdita di generalità, possiamo assumere che tali equazioni siano del tipo $t = \dot{0}$.

Consideriamo ora quindi l'equazione $t = \dot{0}$. Se $\mathbb{C} \models t = \dot{0}$, allora vale $\mathscr{K} \models t = \dot{0}$ per ogni campo (algebricamente chiuso). Questo perché $t$ contiene solamente $\dot{0}, \dot{1}, \dot{-}, \dot{+}, *$, quindi $t$ chiuso denota di fatto in $\mathbb{C}$ sempre un numero intero e in questo preciso caso $t^{\mathbb{C}} = t^{\mathbb{Z}} = 0$. Ma $t$ denoterà $\dot{0}$ in tutti i campi, in quanto qualsiasi assioma che identifichi un certo intero con $\dot{0}$ *solo in alcuni campi* (assiomi delle caratteristiche $p$) è negato in $\mathbb{C}$.

Se invece $\mathbb{C} \not\models t = \dot{0}$, allora $t^{\mathbb{Z}} := z \neq 0$. Il caso interessante è quando anche $z \neq 1$. Allora $\mathbb{Z}_p \models t = \dot{0}$ per $p$ numero primo divisore di $|z|$. Infatti $t^{\mathbb{Z}_p} = Resto(\frac{|t^{\mathbb{Z}}|}{p}) = 0$. Per immersione, $t = \dot{0}$ rimane soddisfatta in tutti i campi di di caratteristica $p$, e questi sono *esattamente* i campi in cui è vera.

Pertanto, per verificare la validità di $\varphi$ in tutti i campi (algebricamente chiusi) è sufficiente verificarne la validità in $\mathbb{C}$ ed in tutti i campi di caratteristica al massimo uguale a

$$p' := \min\{p \mid p \text{ è primo e } p > n\}$$

per il massimo $n$ tale che in $\varphi$ occorre un termine $t$ con $|t^{\mathbb{Z}}| = n$. La procedura di verifica termina quindi in tempo finito.

Abbiamo così dimostrato:

**Teorema 8.10.** *La teoria dei Campi Algebricamente Chiusi è decidibile.*

**Esempio 8.11.** *Consideriamo l'enunciato*

$$\dot{1} \dot{+} \dot{1} \dot{+} \dot{1} = \dot{0} \rightarrow (\dot{1} \dot{+} \dot{1}) \dot{+} (\dot{1} \dot{+} \dot{1}) \dot{+} (\dot{1} \dot{+} \dot{1}) = \dot{0}.$$

*Questo è un teorema della teoria $T$ dei Campi Algebricamente Chiusi come il Teorema di Completezza dimostra. Infatti nei soli campi dove l'antecedente è vero, cioè i campi di caratteristica* 3*, è vero pure il conseguente, giacché* $[(\dot{1} \dot{+} \dot{1}) \dot{+} (\dot{1} \dot{+} \dot{1}) \dot{+} (\dot{1} \dot{+} \dot{1})]^{\mathbb{Z}} = 6$.

*Al contario, l'enunciato*

$$\dot{1} \dotplus \dot{1} \dotplus \dot{1} = \dot{0} \to (\dot{1} \dotplus \dot{1}) \dotplus (\dot{1} \dotplus \dot{1}) \dotplus (\dot{1} \dotplus \dot{1}) \dotplus \dot{1} = \dot{0}.$$

*non sarà un teorema della teoria, come dimostra il Teorema di Validità, perché nei campi dove è vero l'antecedente, ovvero i campi di caratteristica* $3$, *è falso il conseguente, poiché* $[(\dot{1} \dotplus \dot{1}) \dotplus (\dot{1} \dotplus \dot{1}) \dotplus (\dot{1} \dotplus \dot{1}) \dotplus \dot{1}]^{\mathbb{Z}} = 7$.

# Capitolo 9

# Ultrafiltri

## 9.1 Algebre di Boole

Un'algebra di Boole nel linguaggio $\mathscr{L} := \{\dot{0}, \dot{1}, \sqcap, \sqcup\}$ è una struttura $\mathscr{B} := \langle B, \dot{0}, \dot{1}, \sqcap, \sqcup \rangle$, per $B$ insieme non vuoto, $\dot{0}, \dot{1} \in B$, $\sqcap, \sqcup : B \times B \to B$, che soddisfa i seguenti assiomi:

1. $(\forall x)(\forall y)\, x \Diamond y = y \Diamond x$ per $\Diamond \in \{\sqcap, \sqcup\}$ (commutatività)
2. $(\forall x)(\forall y)(\forall z)\, x \Box (y \Diamond z) = (x \Box y) \Diamond (x \Box z)$ per $\Box, \Diamond \in \{\sqcap, \sqcup\}$, $\Box \neq \Diamond$ (distributività)
3. $(\forall x)\, x \sqcap \dot{1} = x$
4. $(\forall x)\, x \sqcup \dot{0} = x$
5. $(\forall x)(\exists y)\, (x \sqcap y = \dot{0} \wedge x \sqcup y = \dot{1})$
6. $\dot{0} \neq \dot{1}$

Si può dimostrare che l'$y$ caratterizzato dall'assioma 5 è univocamente determinato. Siano infatti $y', y''$ siffatti per un determinato $x$. Allora $y' = y' \sqcup \dot{0} = y' \sqcup (x \sqcap y'') = (y' \sqcup x) \sqcap (y' \sqcup y'') = \dot{1} \sqcap (y' \sqcup y'') = y' \sqcup y''$. D'altro canto, analogamente, $y'' = y'' \sqcup \dot{0} = y'' \sqcup (x \sqcap y') = (y'' \sqcup x) \sqcap (y'' \sqcup y') = \dot{1} \sqcap (y'' \sqcup y') = y'' \sqcup y'$. Da questo segue $y' = y''$.

Pertanto possiamo formulare in modo naturale gli assiomi dell'algebra di Boole nel linguaggio $\mathscr{L} := \{\dot{0}, \dot{1}, \dot{-}, \sqcap, \sqcup\}$, dove $\dot{-} : B \to B$ è l'operazione unaria di complementazione che soddisfa l'assioma:

5'. $(\forall x)(x \sqcap \overline{x} = \dot{0} \wedge x \sqcup \overline{x} = \dot{1})$

Si può anche dimostrare che

$$x \sqcap \dot{0} = (x \sqcap \dot{0}) \sqcup \dot{0} = (x \sqcap \dot{0}) \sqcup (x \sqcap \overline{x}) = x \sqcap (\dot{0} \sqcup \overline{x}) = x \sqcap \overline{x} = \dot{0}.$$

Analogamente $x \sqcup \dot{1} = \dot{1}$ (esercizio).

Abbiamo inoltre che $x \sqcap y = x$ sse $x \sqcup y = y$. Sia infatti $x \sqcap y = x$; allora:

$$x \sqcup y = (x \sqcap y) \sqcup y = (x \sqcap y) \sqcup (y \sqcap \dot{1}) = (y \sqcap x) \sqcup (y \sqcap \dot{1}) = y \sqcap (x \sqcup \dot{1}) = y \sqcap \dot{1} = y.$$

Dall'altro lato sia $x \sqcup y = y$; allora:

$$x \sqcap y = x \sqcap (x \sqcup y) = (x \sqcup \dot{0}) \sqcap (x \sqcup y) = x \sqcup (\dot{0} \sqcap y) = x \sqcup \dot{0} = x.$$

Ponendo pertanto $x \leq y \Leftrightarrow_{def} x \sqcap y = x \Leftrightarrow x \sqcup y = y$ abbiamo quindi che $\leq$ è un ordinamento parziale con minimo elemento $\dot{0}$ e massimo elemento $\dot{1}$. Per vedere che $\leq$ è un ordinamento parziale osserviamo che:

- $x \leq x$: infatti $x \sqcap x = x$ (idempotenza; esercizio)
- $x \leq y$ e $y \leq z \Rightarrow x \leq z$: infatti $x \sqcap z = (x \sqcap y) \sqcap z = x \sqcap (y \sqcap z) = x \sqcap y = x$ (per associatività, da dimostrare)
- $x \leq y$ e $y \leq x \Rightarrow x = y$: infatti $x = x \sqcap y = y \sqcap x = y$

Possiamo anche dimostrare che $\sqcap$ è l'operazione che restituisce l'infimo (massimo minorante) dell'insieme di due punti, e $\sqcup$ il supremo (minimo maggiorante). Vediamo il caso di $\sqcup$ (quello di $\sqcap$ è analogo). Per la *legge di assorbimento*, dimostrata sopra *en passant*, si ha $x \sqcap (x \sqcup y) = x$, da cui $x \leq x \sqcup y$. Analogamente, $y \leq x \sqcup y$. Pertanto $x \sqcup y$ è maggiorante di $\{x, y\}$. Vediamo ora anche che è il minimo tra questi. Sia infatti $z$ un maggiorante di $\{x, y\}$. Ciò significa che $x \leq z$ e $y \leq z$, ovvero $x \sqcap z = x$ e $y \sqcap z = y$. Per distributività $(x \sqcup y) \sqcap z = (x \sqcap z) \sqcup (y \sqcap z) = x \sqcup y$, ovvero $x \sqcup y \leq z$.

Infine, si può dimostrare che $\sqcap$ e $\sqcup$ sono monotone, ovvero se $x \leq x'$ e $y \leq y'$ allora $x \diamond y \leq x' \diamond y'$ per $\diamond \in \{\sqcap, \sqcup\}$ (esercizio).

Un classico esempio di algebra di Boole è dato da $\langle \wp(X), \emptyset, X, X \setminus \cdot, \cap, \cup \rangle$ per un qualsiasi insieme $X \neq \emptyset$ fissato. In questo caso l'ordinamento parziale indotto $\leq$ coincide con $\subseteq$.

**Definizione 9.1** (filtri, ultrafiltri). *Sia $\mathscr{B}$ un'algebra di Boole sull'insieme $B$. Allora $\emptyset \neq F \subseteq B$ è un filtro se:*

- $b_1, b_2 \in F \Rightarrow b_1 \sqcap b_2 \in F$ *(chiusura per infimi)*
- $b_1 \in F$ *e* $b_1 \leq b_2 \Rightarrow b_2 \in F$ *(chiusura per sopraelementi).*

*Un filtro $F$ è* proprio *se*

- $\dot{0} \notin F$.

*Un* ultrafiltro *è un filtro proprio tale che*

- $b \in F$ *o* $\overline{b} \in F$ *per ogni* $b \in B$.

Un sottoinsieme $X \subseteq B$ ha la *proprietà dell'intersezione finita (p.i.f.)* se $x_1, ..., x_n \in X \Rightarrow x_1 \sqcap ... \sqcap x_n \neq \dot{0}$. Ovviamente un filtro è proprio sse gode della proprietà dell'intersezione finita.

## 9.2 Dalla compattezza all'estendibilità

**Lemma 9.2.** *Sia $\mathscr{B}$ un'algebra di Boole sull'insieme $B$. Dato $\emptyset \neq X \subseteq B$ con p.i.f., e dati $b_1, ..., b_n \in B$ esiste un filtro proprio $F \supseteq X$ tale che $b_i \in F$ o $\overline{b_i} \in F$ per $1 \leq i \leq n$.*

*Dimostrazione.* Costruiamo il filtro $F$ per induzione:

Passo 0: dato $X$ definiamo

$$F_0 := \{y \in B | (\exists x_1 \in X)...(\exists x_m \in X)\, x_1 \sqcap ... \sqcap x_m \leq y,$$

$$\text{per qualche } m \geq 1\}.$$

Vediamo che $F_0$ è un filtro proprio con $X \subseteq F_0$:

- $F_0$ è non vuoto: per ogni $k \geq 1$ si ha $x_1 \sqcap ... \sqcap x_k \leq \dot{1}$
- $F_0$ è chiusa per infimi: siano $y_1, y_2 \in F_0$. Allora esistono $x_1, ..., x_k$ e $x'_1, ..., x'_l$ in $X$ rispettivamente tali che $x_1 \sqcap ... \sqcap x_k \leq y_1$ e $x'_1 \sqcap ... \sqcap x'_l \leq y_2$. Allora $x_1 \sqcap ... \sqcap x_k \sqcap x'_1 \sqcap .... \sqcap x'_l \leq y_1 \sqcap y_2$ (monotonia)

- $F_0$ è chiuso per sopraelementi: sia $y \in F_0$ con $y \leq z$ e siano $x_1, \dots, x_k \in X$ tali che $x_1 \sqcap \dots \sqcap x_k \leq y$. Allora $x_1 \sqcap \dots \sqcap x_n \leq z$ per transitività di $\leq$

- $\dot{0} \notin F_0$: altrimenti avremmo $x_1 \sqcap \dots \sqcap x_k \leq \dot{0}$ per opportuni $x_1, \dots, x_k \in X$, ovvero $x_1 \sqcap \dots \sqcap x_k = \dot{0}$, il che è impossibile perché $X$ soddisfa la proprietà dell'intersezione finita

- $X \subseteq F_0$: infatti per ogni $x \in B$ e in particolare per ogni $x \in X$ si dà $x \leq x$ (abbiamo quindi un'altra dimostrazione che $F_0$ è non vuoto)

Passo $m+1$ per ipotesi esiste un filtro proprio $F_m \supseteq X$ tale che $b_i \in F_m$ o $\overline{b_i} \in F_m$ per $1 \leq i \leq m$. Dimostriamo che esiste un filtro proprio $F_{m+1} \supseteq F_m$ tale che $b_{m+1} \in F_{m+1}$ o $\overline{b_{m+1}} \in F_{m+1}$. Supponiamo che ciò non si dia. Allora né $F_m \cup \{b_{m+1}\} \subseteq B$ né $F_m \cup \{\overline{b_{m+1}}\} \subseteq B$ hanno la proprietà dell'intersezione finita, perché altrimenti sarebbero estendibili a filtri propri secondo le linee della dimostrazione del passo 0. Quindi esistono $b'_1, \dots b'_k, b''_1, \dots, b''_l \in F_m$ tali che $b'_1 \sqcap \dots \sqcap b'_k \sqcap b_{m+1} = \dot{0}$ e $b''_1 \sqcap \dots \sqcap b''_l \sqcap \overline{b_{m+1}} = \dot{0}$. Poniamo ora $c := b'_1 \sqcap \dots \sqcap b'_k \sqcap b''_1 \sqcap \dots \sqcap b''_l$. Allora $c \in F_m$ e $c \neq \dot{0}$, perché $F_m$ è un filtro proprio. Sappiamo che $c \sqcap b_{m+1} = \dot{0} = c \sqcap \overline{b_{m+1}}$. Ma allora $\dot{0} = (c \sqcap b_{m+1}) \sqcup (c \sqcap \overline{b_{m+1}}) = c \sqcap (b_{m+1} \sqcup \overline{b_{m+1}}) = c \sqcap \dot{1} = c$. Questo è in contraddizione col fatto che $c \neq \dot{0}$.

□

**Teorema 9.3.** *Sia $\mathscr{B}$ un'algebra di Boole sull'insieme $B$. Allora $\emptyset \neq X \subseteq B$ con p.i.f. è estendibile ad un ultrafiltro.*

*Dimostrazione.* Consideriamo

$$Th(\mathscr{B}) := \{\varphi \mid \varphi \text{ è } \mathscr{L}_B\text{-formula chiusa e } \mathscr{B} \models \varphi\}$$

dove $\mathscr{L}_B$ estende $\mathscr{L}$ possedendo una costante per ciascun $b \in B$ (per semplicità, non distingueremo tali costanti dalle loro denotazioni). Consideriamo ora il predicato $U(x)$ letto intuitivamente come "$x \in F$ per $F \supseteq X$ ultrafiltro", e che è assiomatizzato dal seguente insieme $\Gamma$ di $\mathscr{L}_B$-formule:

- $U(b)$ per ogni $b \in X$

- $U(b_1) \wedge U(b_2) \to U(b_1 \sqcap b_2)$ per ogni $b_1, b_2 \in B$
- $U(b_1) \to U(b_2)$ per ogni $b_1, b_2 \in B$ tali che $b_1 \leq b_2$
- $\neg U(\dot{0})$
- $\neg U(b) \to U(\overline{b})$ per ogni $b \in B$

Consideriamo ora qualsiasi sottoinsieme finito $\Gamma_0 \subseteq Th(\mathscr{B}) \cup \Gamma$ e sia $\mathscr{B}'$ l'estensione di $\mathscr{B}$ che interpreta $U$ come il filtro proprio $F$ definito come nel Lemma 9.2 relativamente agli individui $b_1, ..., b_n$ che compaiono in $\Gamma_0$ (senza perdita di generalità possiamo infatti assumere che $\Gamma_0 \cap \Gamma \neq \emptyset$). Abbiamo $\mathscr{B}' \models \Gamma_0$.

Allora $Th(\mathscr{B}) \cup \Gamma$ è finitamente soddisfacibile, e per il Teorema di Compattezza (versione generale, per linguaggi infiniti arbitrariamente grandi) è anche soddisfacibile. Pertanto esiste un ultrafitro che estende $X$. □

## 9.3 *Dall'estendibilità alla compattezza*

Abbiamo quindi visto che il Teorema di Compattezza implica l'estendibilità degli insiemi con p.i.f. ad ultrafiltri. Vediamo ora che, al contrario, tale estendibilità implica il Teorema di Compattezza (Poizat [2012]).

**Teorema 9.4.** *Un filtro proprio è un ultrafiltro sse è massimale per inclusione.*

*Dimostrazione.* $\Rightarrow$) Sia per assurdo $F$ un ultrafiltro su un algebra di Boole $\mathscr{B}$ con dominio $B$ che non sia non massimale come filtro proprio, ovvero $F \subset F'$ per $F'$ filtro proprio su $B$. Allora esiste $b \in B$ tale che $b \in F'$ e $b \notin F$. Ma $F$ è ultrafiltro, e quindi $\overline{b} \in F$. Poiché $F \subset F'$, abbiamo $\overline{b} \in F'$, da cui $b \sqcap \overline{b} = \dot{0} \in F'$, il che è assurdo perché $F'$ è un filtro proprio.

$\Leftarrow$) Sia $F$ massimale per inclusione tra i filtri propri sull'algebra di Boole $\mathscr{B}$ con dominio $B$. Sia $b^* \notin F$ (si noti che almeno un tal $b^*$ esiste, essendo $F$ proprio). Mostriamo allora che $\overline{b^*} \in F$. Definiamo $F' := \{a \in B \mid (\exists c \in F)\, c \sqcap b^* \leq a\}$, e osserviamo che se $a \in F$, allora $a \sqcap b^* \leq a$, quindi $a \in F'$, da cui $F \subseteq F'$. Per ipotesi $b^* \notin F$, in compenso $b^* = \dot{1} \sqcap b^*$, da cui $b^* \in F'$. Da questo segue che $F \subset F'$, e quindi, per massimalità di $F$, abbiamo che $F'$ non può essere un filtro proprio. Tuttavia $F'$ soddisfa:

- la chiusura rispetto agli infimi: siano $a, a' \in F'$. Allora esistono $c, c' \in F$ tali che $c \sqcap b^* \leq a$, $c' \sqcap b^* \leq a'$, da cui $(c \sqcap c') \sqcap b^* = (c \sqcap c') \sqcap (b^* \sqcap b^*) = (c \sqcap b^*) \sqcap (c' \sqcap b^*) \leq a \sqcap a'$ (monotonia); inoltre $c \sqcap c' \in F$ perché $F$ è un filtro, da cui $a \sqcap a' \in F'$

- chiusura per sopraelementi: sia $a \in F'$ e $a \leq a'$. Allora esiste $c \in F$ tale che $c \sqcap b^* \leq a$, da cui $c \sqcap b^* \leq a'$, e quindi $a' \in F'$.

Pertanto $F'$ è un filtro. Per questa ragione, non potendo essere un filtro proprio, si dovrà avere $\dot{0} \in F'$. Perciò esiste un $c \in F$ tale che $c \sqcap b^* \leq \dot{0}$, cioè $c \sqcap b^* = \dot{0}$, da cui $(c \sqcap b^*) \sqcup \overline{b^*} = \dot{0} \sqcup \overline{b^*} = \overline{b^*}$, ma anche $(c \sqcap b^*) \sqcup \overline{b^*} = (c \sqcup \overline{b^*}) \sqcap (b^* \sqcup \overline{b^*}) = (c \sqcup \overline{b^*}) \sqcap \dot{1} = c \sqcup \overline{b^*}$; da questo segue $c \sqcup \overline{b^*} = \overline{b^*}$, ovvero $c \leq \overline{b^*}$; essendo $F$ un filtro, e quindi chiuso per sopraelementi, concludiamo $\overline{b^*} \in F$. □

Il prossimo teorema dimostra che in un'algebra di Boole ogni insieme con la p.i.f. può essere esteso ad un ultrafiltro, cioè non è altro che il già visto Teorema 9.3. Tuttavia questa volta non utilizzeremo il Teorema di Compattezza per dimostrarlo.

**Teorema 9.5.** *Sia $\mathscr{B}$ un'algebra di Boole sull'insieme $B$. Allora $\emptyset \neq X \subseteq B$ con p.i.f. è estendibile ad un ultrafiltro.*

*Dimostrazione.* Sappiamo già per il Lemma 9.2 che $X$ è estendibile ad un filtro proprio $F$. Poniamo ora

$$\Omega_F := \left\{ F' \subseteq B \mid F' \text{ è un filtro proprio e } F \subseteq F' \right\}.$$

Ovviamente $\Omega_F$ è parzialmente ordinato da $\subseteq$. Sia quindi $\mathscr{C} \subseteq \Omega_F$ una catena (cioè un sottoinsieme totalmente ordinato), e vediamo che $\bigcup \mathscr{C}$ è un maggiorante di $\mathscr{C}$ in $\Omega_F$, ovvero (i) $\bigcup \mathscr{C} \in \Omega_F$ e (ii) se $F' \in \mathscr{C}$ allora $F' \subseteq \bigcup \mathscr{C}$. La proprietà (ii) è immediata per la definizione di $\bigcup$. Per la (i) osserviamo innanzitutto che, poiché $F \subseteq F'$ per qualsiasi $F' \in \mathscr{C}$, si ha banalmente $F \subseteq \bigcup \mathscr{C}$, e pertanto, per mostrare che $\bigcup \mathscr{C} \in \Omega_F$, resta solo da dimostrare che $\bigcup \mathscr{C}$ è un filtro proprio:

- $\dot{0} \notin \bigcup \mathscr{C}$: infatti $\dot{0} \notin F'$ per ogni $F' \in \mathscr{C}$

- chiusura per infimi: siano $b, b' \in \bigcup\mathscr{C}$, allora $b \in F'$ e $b' \in F''$ per $F', F'' \in \mathscr{C}$, ed essendo $\mathscr{C}$ totalmente ordinato possiamo supporre che $F' \subseteq F''$, e quindi $b, b' \in F''$. Questo significa che anche $b \sqcap b' \in F'' \in \mathscr{C}$, poiché $F''$ è un filtro, da cui $b \sqcap b' \in \bigcup\mathscr{C}$

- chiusura per sopraelementi: sia $b \in \bigcup\mathscr{C}$ e $b \leq b'$. Allora $b \in F'$ per qualche $F' \in \mathscr{C}$, da cui anche $b' \in F'$ poiché $F'$ è un filtro. Pertanto $b' \in \bigcup\mathscr{C}$.

Quindi ogni catena in $\Omega_F$ ha un maggiorante in $\Omega_F$, e quindi per il Lemma di Zorn $\Omega_F$ ammette un elemento massimale $U$, il quale sarà un filtro proprio che estende $F$, e per il Teorema 9.4 sarà anche un ultrafiltro. □

Siano dati un insieme di indici $I$ ed un ultrafiltro $U$ su $\wp(I)$. Per semplicità, nel seguito diremo che $U$ è un *ultrafiltro su* $I$.

**Definizione 9.6** (equivalenza ultrafiltrica). *Sia dato un insieme di indici $I$. Data una famiglia di insiemi $\{A_i\}_{i\in I}$ ed $U$ ultrafiltro su $I$ definiamo sull'insieme $\Pi_{i\in I} A_i$ (l'insieme delle funzioni di scelta sugli $A_i$) la relazione $(a_i)_{i\in I} \sim_U (a'_i)_{i\in I} \Leftrightarrow_{def} \left\{ i \in I \mid a_i = a'_i \right\} \in U$ (in questo caso diciamo che $(a_i)_{i\in I}$* è quasi sempre *uguale a $(a'_i)_{i\in I}$). Ove non occorra ambiguità possiamo scrivere semplicemente $\sim$ invece di $\sim_U$.*

**Proposizione 9.7.** *$\sim_U$ è relazione di equivalenza.*

*Dimostrazione.*

- Riflessività: perché $I \in U$

- Simmetria: per simmetria di $=$

- Transitività: per chiusura per intersezioni (infimi) e per soprainsiemi (sopraelementi) di $U$: per $(a_i)_i, (a'_i)_i, (a''_i)_i \in \Pi_i A_i$ si osservi che

$$\left\{ i \in I \mid a_i = a''_i \right\} \supseteq \left\{ i \in I \mid a_i = a'_i = a''_i \right\} =$$

$$= \left\{ i \in I \mid a_i = a'_i \right\} \cap \left\{ i \in I \mid a'_i = a''_i \right\}$$

□

**Definizione 9.8** (ultraprodotto)**.** *Siano dati un insieme di indici $I$, un ultrafiltro $U$ su $I$, ed $\mathscr{L}$-strutture $\mathscr{A}_i$ con domini $A_i$ rispettivamente per $i \in I$. Sia allora $\Pi_U \mathscr{A}_i$ la $\mathscr{L}$-stuttura definita nel seguente modo:*

- *il dominio di $\Pi_U \mathscr{A}_i$ è $\Pi_U A_i := (\Pi_{i\in I} A_i)/\sim_U$*
- *per ogni simbolo di costante $c$: $c^{\Pi_U \mathscr{A}_i} := [(c^{\mathscr{A}_i})_i]$*
- *per ogni simbolo di funzione $n$-ario $f$: $f^{\Pi_U \mathscr{A}_i}([(a_i^1)_i], \dots, [(a_i^n)_i]) = [(a_i)_i]$ sse $\left\{ i \in I \mid f^{\mathscr{A}_i}(a_i^1, \dots, a_i^n) = a_i \right\} \in U$*
- *per ogni simbolo $n$-ario di predicato $P$: $([(a_i^1)_i], \dots, [(a_i^n)_i]) \in P^{\Pi_U \mathscr{A}_i}$ sse $\left\{ i \in I \mid (a_i^1, \dots, a_i^n) \in P^{\mathscr{A}_i} \right\} \in U$*

Si osservi che la definizione di $\Pi_U \mathscr{A}_i$ è ben posta:

- costanti: ovvio
- sia $(a_i^j)_i \sim (a_i'^j)_i$ per $1 \le j \le n$ e sia $f^{\Pi_U \mathscr{A}_i}([(a_i'^1)_i], \dots, [(a_i'^n)_i]) = [(a_i')_i]$ (ovvero $\left\{ i \mid f^{\mathscr{A}_i}(a_i'^1, \dots, a_i'^n) = a_i' \right\} \in U$); allora

  $$\left\{ i \mid a_i = a_i' \right\} \supseteq \bigcap_{1\le j\le n} \left\{ i \mid a_i^j = a_i'^j \right\} \cap \left\{ i \mid f^{\mathscr{A}_i}(a_i^1, \dots, a_i^n) = a_i \right\} \in U$$

  da cui $(a_i)_i \sim (a_i')_i$
- sia $(a_i^j)_i \sim (a_i'^j)_i$ per $1 \le j \le n$; allora

  $$\left\{ i \mid (a_i'^1, \dots, a_i'^n) \in P^{\mathscr{A}_i} \right\} \supseteq$$

  $$\supseteq \bigcap_{1\le j\le n} \left\{ i \mid a_i^j = a_i'^j \right\} \cap \left\{ i \mid (a_i^1, \dots, a_i^n) \in P^{\mathscr{A}_i} \right\} \in U$$

  da cui $\left\{ i \mid (a_i'^1, \dots, a_i'^n) \in P^{\mathscr{A}_i} \right\} \in U$ e quindi $([(a_i'^1)_i], \dots, [(a_i'^n)_i]) \in P^{\Pi_U \mathscr{A}_i}$

**Teorema 9.9** (Teorema di Łos)**.** *Sia $\varphi(x_1, \dots, x_n)$ una $\mathscr{L}$-formula e siano $[(a_i^1)_i], \dots, [(a_i^n)_i] \in \Pi_U A_i$. Allora:*

$$\Pi_U \mathscr{A}_i \models \varphi[[(a_i^1)_i], \dots, [(a_i^n)_i]] \iff \left\{ i \in I \mid \mathscr{A}_i \models \varphi[a_i^1, \dots, a_i^n] \right\} \in U.$$

*Dimostrazione.* La dimostrazione è per induzione sulla costruzione di $\varphi$ e ne analizzeremo i casi fondamentali. Per prima cosa osserviamo che per induzione sulla costruzione degli $\mathscr{L}$-termini $t(x_1, \ldots, x_n)$ si può dimostrare che per ogni $[(a_i^1)_i], \ldots, [(a_i^n)_i], [(a_i)_i] \in \Pi_U A_i$ si ha

$$t^{\Pi_U \mathscr{A}_i}([(a_i^1)_i], \ldots, [(a_i^n)_i]) = [(a_i)_i] \Longleftrightarrow \left\{ i \mid t^{\mathscr{A}_i}(a_i^1 \ldots, a_i^n) = a_i \right\} \in U \qquad (\clubsuit)$$

Allora:

- $\varphi(x_1, \ldots, x_n) :\equiv P(t_1(x_1, \ldots, x_n), \ldots, t_m(x_1, \ldots, x_n))$:
  sia $t_j^{\Pi_U \mathscr{A}_i}([(a_i^1)_i], \ldots, [(a_i^n)_i]) = [(a_i'^j)_i]$ per $1 \leq j \leq m$; allora:

$$\Pi_U \mathscr{A}_i \models P(t_1(x_1, \ldots, x_n), \ldots, t_m(x_1, \ldots, x_n))[[(a_i^1)_i], \ldots, [(a_i^n)_i]]$$

sse

$$(t_1^{\Pi_U \mathscr{A}_i}([(a_i^1)_i], \ldots, [(a_i^n)_i]), \ldots, t_m^{\Pi_U \mathscr{A}_i}([(a_i^1)_i], \ldots, [(a_i^n)_i])) \in P^{\Pi_U \mathscr{A}_i}$$

sse

$$([(a_i'^1)_i], \ldots, [(a_i'^m)_i]) \in P^{\Pi_U \mathscr{A}_i}$$

sse

$$\left\{ i \mid (a_i'^1, \ldots, a_i'^m) \in P^{\mathscr{A}_i} \right\} \in U$$

sse

$$\left\{ i \mid (a_i'^1, \ldots, a_i'^m) \in P^{\mathscr{A}_i} \right\} \cap \bigcap_{1 \leq j \leq m} \left\{ i \mid a_i'^j = t_j^{\mathscr{A}_i}(a_i^1, \ldots, a_i^n) \right\} \in U (\text{per } (\clubsuit))$$

sse

$$\left\{ i \mid (t_1^{\mathscr{A}_i}(a_i^1, \ldots, a_i^n), \ldots, t_m^{\mathscr{A}_i}(a_i^1, \ldots, a_i^n)) \in P^{\mathscr{A}_i} \right\} \in U$$

sse

$$\left\{ i \mid \mathscr{A}_i \models P(t_1(x_1, \ldots, x_n), \ldots, t_m(x_1, \ldots, x_n))[a_i^1, \ldots, a_i^n] \right\} \in U$$

- $\varphi(x_1, ..., x_n) :\equiv t_1(x_1, ..., x_n) = t_m(x_1, ..., x_n)$: analogo
- $\varphi(x_1, ..., x_n) :\equiv \neg\psi(x_1, ..., x_n)$: allora:

$$\Pi_U \mathscr{A}_i \models \neg\psi(x_1, ..., x_n)[[(a_i^1)_i], ..., [(a_i^n)_i]]$$

sse

$$\Pi_U \mathscr{A}_i \not\models \psi[[(a_i^1)_i], ..., [(a_i^n)_i]]$$

sse

$$\{\, i \mid \mathscr{A}_i \not\models \psi[a_i^1, ..., a_i^n]\,\} \notin U \text{ (per IH)}$$

sse

$$I \setminus \{\, i \mid \mathscr{A}_i \not\models \psi[a_i^1, ..., a_i^n]\,\} \in U$$

sse

$$\{\, i \mid \mathscr{A}_i \models \neg\psi(x_1, ..., x_n)[a_i^1, ..., a_i^n]\,\} \in U$$

- $\varphi(x_1, ..., x_n) \equiv \psi(x_1, ..., x_n) \wedge \chi(x_1, ..., x_n)$: allora:

$$\Pi_U \mathscr{A}_i \models \psi(x_1, ..., x_n) \wedge \chi(x_1, ..., x_n)[[(a_i^1)_i], ..., ([(a_i^n)_i]]$$

sse

$$\Pi_U \mathscr{A}_i \models \psi[[(a_i^1)_i], ..., [(a_i^n)_i]] \text{ e } \Pi_U \mathscr{A}_i \models \chi[[(a_i^1)_i], ..., [(a_i^n)_i]]$$

sse

$$\{\, i \mid \mathscr{A}_i \models \psi[a_i^1, ..., a_i^n]\,\} \in U \text{ e } \{\, i \mid \mathscr{A}_i \models \chi[a_i^1, ..., a_i^n]\,\} \in U \text{(per IH)}$$

sse

$$\{\, i \mid \mathscr{A}_i \models \psi[a_i^1, ..., a_i^n]\,\} \cap \{\, i \mid \mathscr{A}_i \models \chi[a_i^1, ..., a_i^n]\,\} \in U$$

sse

$$\{\, i \mid \mathscr{A}_i \models \psi(x_1, \dots, x_n) \wedge \chi(x_1, \dots, x_n)[a_i^1, \dots, a_i^n] \,\} \in U$$

- $\varphi(x_1, \dots, x_n) :\equiv (\exists x)\psi(x, x_1, \dots, x_n)$:

  $\Rightarrow$) valga $\Pi_U \mathscr{A}_i \models (\exists x)\psi(x, x_1, \dots, x_n)[[(a_i^1)_i], \dots, ([(a_i^n)_i]]$. Allora esiste $[(a_i)_i] \in \Pi_U A_i$ tale che

  $$\Pi_U \mathscr{A}_i \models \psi[[(a_i)_i], [(a_i^1)_i], \dots, ([(a_i^n)_i]],$$

  e quindi, per ipotesi d'induzione, $\{\, i \mid \mathscr{A}_i \models \psi[a_i, a_i^1, \dots, a_i^n] \,\} \in U$, da cui segue $\{\, i \mid \mathscr{A}_i \models (\exists x)\psi(x, x_1, \dots, x_n)[a_i^1, \dots, a_i^n] \,\} \in U$.

  $\Leftarrow$) valga $J := \{\, i \mid \mathscr{A}_i \models (\exists x)\psi(x, x_1, \dots, x_n)[a_i^1, \dots, a_i^n] \,\} \in U$. Allora per ogni $i \in J$ selezioniamo un $a_i \in A_i$ tale che

  $$\mathscr{A}_i \models \psi[a_i, a_i^1, \dots, a_i^n].$$

  Inoltre, per ogni $i \notin J$, selezioniamo un individuo $a_i' \in A_i$. Consideriamo ora l'elemento $[(a_i'')_i]$ tale che $a_i'' = a_i$ per $i \in J$ e $a_i'' = a_i'$ per $i \notin J$. Allora

  $$\{\, i \mid \mathscr{A}_i \models \psi[a_i'', a_i^1, \dots, a_i^n] \,\} = \{\, i \mid \mathscr{A}_i \models \psi[a_i, a_i^1, \dots, a_i^n] \,\} \in U;$$

  pertanto per ipotesi d'induzione

  $$\Pi_U \mathscr{A}_i \models \psi[[(a_i'')_i], [(a_i^1)_i], \dots, [(a_i^n)_i]],$$

  da cui $\Pi_U \mathscr{A}_i \models (\exists x)\psi(x, x_1, \dots, x_n)[(a_i^1)_i], \dots, [(a_i^n)_i]]$.

$\square$

Sia ora $\mathscr{T}$ l'insieme delle $\mathscr{L}$-teorie complete e soddisfacibili, con topologia $\tau$ generata dagli insiemi $O_\varphi := \{\, T \in \mathscr{T} \mid \varphi \in T \,\}$ per ogni $\mathscr{L}$-formula $\varphi$ chiusa. Si osservi che gli $O_\varphi$ sono anche chiusi, in quanto $\mathscr{T} \setminus O_\varphi = O_{\neg\varphi}$. Pertanto gli elementi della base sono dei *clopen*, e non solo, i loro complementari sono altri elementi della base.

**Teorema 9.10** (Teorema di Compattezza (Topologica))**.** *$\mathscr{T}$ è compatto.*

*Dimostrazione.* Sia $\{A_i\}_{i \in I}$ una famiglia di insiemi chiusi in $\mathscr{T}$ aventi la p.i.f. Poiché gli aperti della base sono dei clopen aventi per complementari altri insiemi della base, ciascun insieme chiuso di $\mathscr{T}$ è un'intersezione di elementi della base (esercizio). Pertanto, senza perdita di generalità, possiamo considerare solo famiglie di chiusi del tipo $\{O_{\varphi_i}\}_{i\in I}$ (esercizio). Prendiamo quindi una tale famiglia con la p.i.f. Per il Teorema 9.5 tale famiglia si può estendere ad un ultrafiltro $U$ sulla base di $\mathscr{T}$; tale base è infatti un'algebra di Boole (esercizio, osservando che $O_\varphi \cap O_\psi = O_{\varphi\wedge\psi}$ e $O_\varphi \cup O_\psi = O_{\varphi\vee\psi}$). Poiché $\mathscr{T}$ contiene solo teorie soddisfacibili, ciascuna $T$ ha un $\mathscr{L}$-modello $\mathscr{M}_T$. Consideriamo ora la teoria

$$T' := Th(\Pi_U \mathscr{M}_T) = \left\{\varphi \mid \varphi \text{ è } \mathscr{L}\text{-formula chiusa e } \Pi_U \mathscr{M}_T \models \varphi\right\}.$$

Si osservi che $T'$ è soddisfacibile e completa e quindi $T' \in \mathscr{T}$. Abbiamo inoltre $\Pi_U \mathscr{M}_T \models \varphi \Leftrightarrow \mathscr{M}_{T'} \models \varphi$ (esercizio). Sappiamo dal Teorema di Łos 9.9 che $\Pi_U \mathscr{M}_T \models \varphi \Leftrightarrow \left\{T \in \mathscr{T} \mid \mathscr{M}_T \models \varphi\right\} \in U$, ma, poiché stiamo considerando solo teorie complete, $\left\{T \in \mathscr{T} \mid \mathscr{M}_T \models \varphi\right\} = \left\{T \in \mathscr{T} \mid \varphi \in T\right\} = O_\varphi$. Quindi

$$O_\varphi \in U \Longleftrightarrow \Pi_U \mathscr{M}_T \models \varphi \Longleftrightarrow \mathscr{M}_{T'} \models \varphi \Longleftrightarrow \varphi \in T' \Longleftrightarrow T' \in O_\varphi \qquad (*)$$

Ora, siccome $U$ estende $\{O_{\varphi_i}\}_{i\in I}$, abbiamo che $O_{\varphi_i} \in U$ per ogni $i \in I$, da cui per $(*)$ $T' \in O_{\varphi_i}$ per ogni $i \in I$. Ne segue che $\bigcap_{i\in I} O_{\varphi_i} \neq \emptyset$, e pertanto $\mathscr{T}$ è compatto. □

**Corollario 9.11** (Teorema di Compattezza (Logica))**.** *Sia* $\Gamma$ *un insieme di* $\mathscr{L}$*-formule chiuse. Allora* $\Gamma$ *è soddisfacibile sse è finitamente soddisfacibile.*

*Dimostrazione.* $\Rightarrow$) Ovvio.

$\Leftarrow$) Sia dato un qualsiasi sottoinsieme finito $\Gamma' \subseteq \Gamma$. Poiché per ipotesi $\Gamma'$ è soddisfacibile, ha un modello $\mathscr{M}'$. Per tale modello, vale che $\varphi \in Th(\mathscr{M}')$ per ogni $\varphi \in \Gamma'$, e quindi $Th(\mathscr{M}') \in O_\varphi$ per ogni $\varphi \in \Gamma'$, il che significa $\bigcap_{\varphi\in\Gamma'} O_\varphi \neq \emptyset$. Da questo concludiamo che la famiglia $\{O_\varphi\}_{\varphi\in\Gamma}$ ha la p.i.f.. Per il Teorema di Compattezza 9.10, $\bigcap_{\varphi\in\Gamma} O_\varphi \neq \emptyset$. Sia pertanto $T$ un elemento di quest'intersezione. Poiché $T$ è una teoria soddisfacibile, essa ha un modello $\mathscr{M}$. Inoltre $T \supseteq \Gamma$, da cui $\mathscr{M} \models \Gamma$. □

# Capitolo 10

# Analisi Non Standard

## *10.1 Gli iperreali*

**Motivazioni:**[1] *calcolo con gli infinitesimi nel XVII secolo:* sia dato un grave lasciato in caduta libera. Esso cadendo percorre un tragitto spaziale misurato in metri che è funzione del tempo: $s(t) = 4,9t^2\ m$. La velocità media di caduta sarà $\frac{\Delta s}{\Delta t}(t_1, t_2) = \frac{s(t_2)-s(t_1)}{t_2-t_1}\ \frac{m}{s}$.

Ma qual è la velocità di caduta *istantanea*?

Idea! Consideriamo intervalli $\Delta s$ e $\Delta t$ *infinitesimi*, ovvero piccolissimi ma non nulli, ovvero più piccoli di qualsiasi numero reale positivo ma più grandi di 0 (!), indicati con $ds$ e $dt$. La velocità istantanea sarà $\frac{ds}{dt}(t)$.

Calcoliamo ad esempio la velocità istantanea all'istante $t_1 := 1$. A tal fine consideriamo un istante temporale $t_1 + dt = 1 + dt$ "infinitamente" vicino a $t_1$ eppure non coincidente con esso. Allora:

$$s(t_1) = 4,9 \cdot 1^2 = 4,9\ m$$

$$s(t_1+dt) = 4,9\cdot(1+dt)^2 = 4,9(1+2dt+(dt)^2) = 4,9+9,8dt+4,9(dt)^2 m$$

$$ds(t_1) = s(t_1 + dt) - s(t_1) = 9,8dt + 4,9(dt)^2\ m$$

$$\frac{ds}{dt}(t_1) = \frac{9,8\,\cancel{d}t + 4,9(dt)^{\cancel{2}}}{\cancel{d}t} = 9,8 + 4,9dt\ m/s$$

[1] La seguente discussione storico-concettuale ha meramente un valore didattico. Per la ricostruzione storica "senza pretese" ci riferiamo a Davis & Hersh [1972. Ristampa in "Logica", Le Scienze Quaderni 60, a cura di C. Mangione, 1991].

Tuttavia $t_1 = 1$ è un istante di tempo reale, e quindi ci aspettiamo una velocità istantanea reale, non contenente infinitesimi che nella realtà fisica non esistono!

Idea! $dt$ è una quantità molto piccola, quindi la possiamo trascurare (anche in un prodotto con un reale), e pertanto... cancellare! Otteniamo così:

$$\frac{ds}{dt}(1) = 9,8\ m/s.$$

*Pragmatica:* questo risultato è reale ed esatto, ed è quello che otteniamo calcolando la derivata $s'$ di $s$ rispetto a $t$ ed applicandola a $t = 1$ secondo le attuali regole dell'analisi basate sul concetto di limite.

*Logica:* le regole di questo calcolo sono contraddittorie: quando ci fa comodo consideriamo $dt \neq 0$: così non fosse avremmo $t_1 + dt = t_1$, $ds = dt = 0$ (e $\frac{ds}{dt}$ manco sarebbe determinato!); quando ci fa comodo consideriamo $dt = 0$, e infatti alla fine lo eliminiamo.

*Critica di George Berkeley:* "Che sono queste flussioni[2]"? Le velocità d'incrementi evanescenti. E che sono questi stessi incrementi evanescenti? Non sono né quantità finite[3], né quantità infinitamente piccole, e neppure nulla. Perché non chiamarle fantasmi di quantità svanite?" (L'Analista, 1734)

*Soluzione Weiertraßiana:* al bando gl'infinitesimi, usiamo i limiti, le approssimazioni mediante $\varepsilon$ e $\delta$, e calcoliamo così le derivate.

*Logica:* ..... eppure si dà il Teorema di Compattezza! Dato un qualsiasi sottoinsieme finito $I \subseteq \mathbb{N}$, abbiamo

$$\mathbb{R} \models \left\{ (\exists x)\, 0 < x < \frac{1}{n} \mid n \in I \right\}.$$

Pertanto.... deve esistere una struttura $\mathbb{R}^*$ tale che

$$\mathbb{R}^* \models \left\{ (\exists x)\, 0 < x < \frac{1}{n} \mid n \in \mathbb{N} \right\}!$$

*Pragmatica:* ragionare in termini di infinitesimi pare più semplice che ragionare in termini di limiti e approssimazioni con $\varepsilon, \delta$.

*Ontologia:* ammettendo anche gli infinitesimi otteniamo un universo numerico più ricco.

---

[2]O grandezze istantanee, quantità infinitesime.

[3]Nel senso di quantità reali (diverse da 0), quindi non "*infinitamente piccole*".

*Estetica:* l'analisi *infinitesimale* è molto bella!

Problema: come costruiamo $\mathbb{R}^*$? Idea: usiamo gli ultrafiltri!

In Analisi Non Standard (Goldblatt [1998]) considereremo un linguaggio $\mathscr{L}_\mathbb{R}$ per $\mathbb{R}$ che contenga costanti per tutti i numeri reali, simboli per tutte le funzioni (parziali) (!) e per tutti i predicati in $\mathbb{R}$ (in corrispondenza biunivoca). Notazionalmente, in analogia con la pratica matematica, non distingueremo tra le costanti numeriche, i simboli di funzione, i simboli di predicato da un lato e le loro denotazioni dall'altro. La cosa si estenderà in modo automatico a tutti gli $\mathbb{R}$-termini chiusi. Inoltre si concedono funzioni parziali perché si ammette che non tutti i termini siano necessariamente interpretati: ad esempio, $\frac{0}{0}, \tan(\frac{\pi}{2})$ non saranno interpretati; in questo modo abbiamo una semantica più vicina all'ordinaria pratica matematica. Diciamo pertanto che un termine $t$ chiuso è *definito* se $t$ ha un'interpretazione in $\mathbb{R}$:

- per ogni costante $r$, $r$ è definito
- per ogni simbolo di funzione $n$-ario $f$ e per $t_1, ..., t_n$ $\mathbb{R}$-termini chiusi, $f^n(t_1, ..., t_n)$ è definito se $t_i$ è definito per $1 \leq i \leq n$ e $(t_1, ..., t_n) \in \mathrm{dom}(f)$.

Per estensione, diciamo che una formula chiusa è *definita* se tutti i termini presenti in essa sono definiti.

Sulla base di $\mathscr{L}_\mathbb{R}$, apportiamo anche una leggera variazione al linguaggio del primordine, considerando che tutti i quantificatori siano limitati, ovvero del tipo $(\forall x \in A)$, $(\exists x \in A)$, per $A \subseteq \mathbb{R}$.

La definizione di verità in $\mathbb{R}$ delle formule diventa pertanto la seguente:

- $\mathbb{R} \models P(t_1(x_1, ...x_n), ..., t_m(x_1, ..., x_n))[r_1, ..., r_n]$
  qualora $P(t_1(r_1/x_1, ..., r_n/x_n), ..., t_m(r_1/x_1, ..., r_n/x_n))$ sia definita e $(t_1(r_1, ..., r_n), ..., t_m(r_1, ..., r_n)) \in P$
- $\mathbb{R} \models \neg\psi(x_1, ..., x_n)[r_1, ..., r_n]$ qualora $\mathbb{R} \not\models \psi(x_1, ..., x_n)[r_1, ..., r_n]$, per $\psi(r_1/x_1, ..., r_n/x_n)$ definita
- $\mathbb{R} \models \psi(x_1, ..., x_n) \wedge \chi(x_1, ..., x_n)[r_1, ...., r_n]$
  qualora $\mathbb{R} \models \psi(x_1, ..., x_n)[r_1, ..., r_n]$ e $\mathbb{R} \models \chi(x_1, ..., x_n)[r_1, ..., r_n]$, per $\psi(r_1/x_1, ..., r_n/x_n)$ e $\chi(r_1/x_1, ..., r_n/x_n)$ definite

- $\mathbb{R} \models \psi(x_1, \ldots, x_n) \vee \chi(x_1, \ldots, x_n)[r_1, \ldots., r_n]$
  qualora $\mathbb{R} \models \psi(x_1, \ldots, x_n)[r_1, \ldots, r_n]$ o $\mathbb{R} \models \chi(x_1, \ldots, x_n)[r_1, \ldots, r_n]$, per $\psi(r_1/x_1, \ldots, r_n/x_n)$ e $\chi(r_1/x_1, \ldots, r_n/x_n)$ definite

- $\mathbb{R} \models \psi(x_1, \ldots, x_n) \to \chi(x_1, \ldots, x_n)[r_1, \ldots., r_n]$
  qualora $\mathbb{R} \not\models \psi(x_1, \ldots, x_n)[r_1, \ldots, r_n]$ o $\mathbb{R} \models \chi(x_1, \ldots, x_n)[r_1, \ldots, r_n]$, per $\psi(r_1/x_1, \ldots, r_n/x_n)$ e $\chi(r_1/x_1, \ldots, r_n/x_n)$ definite

- $\mathbb{R} \models \psi(x_1, \ldots, x_n) \leftrightarrow \chi(x_1, \ldots, x_n)[r_1, \ldots., r_n]$
  qualora $\mathbb{R} \models \psi(x_1, \ldots, x_n)[r_1, \ldots, r_n]$ sse $\mathbb{R} \models \chi(x_1, \ldots, x_n)[r_1, \ldots, r_n]$, per $\psi(r_1/x_1, \ldots, r_n/x_n)$ e $\chi(r_1/x_1, \ldots, r_n/x_n)$ definite

- $\mathbb{R} \models (\forall x \in P)\, \psi(x, x_1, \ldots, x_n)[r_1, \ldots, r_n]$
  qualora $\psi(r/x, r_1/x_1, \ldots, r_n/x_n)$ sia definita
  e $\mathbb{R} \models \psi(x, x_1, \ldots, x_n)[r, r_1, \ldots, r_n]$ per ogni $r \in P$

- $\mathbb{R} \models (\exists x \in P)\, \psi(x, x_1, \ldots, x_n)[r_1, \ldots, r_n]$
  qualora $\psi(r/x, r_1/x_1, \ldots, r_n/x_n)$ sia definita
  e $\mathbb{R} \models \psi(x, x_1, \ldots, x_n)[r, r_1, \ldots, r_n]$ per qualche $r \in P$

Fissiamo ora un ultrafiltro $U$ su $\mathbb{N}$ che sia *non principale*, ovvero che non contenga sottoinsiemi finiti di $\mathbb{N}$. Allora il dominio di $\mathbb{R}^*$ è definito come l'insieme $\Pi_U \mathbb{R}$.

Dato un numero reale $r$, una funzione $f$ (parziale) su $\mathbb{R}$, un predicato (relazione) $P$ su $\mathbb{R}$, definiamo le loro controparti $r^*$, $f^*$, $P^*$ in $\mathbb{R}^*$ come previsto dalla definizione di ultraprodotto (a parte il naturale adattamento alle funzioni parziali, pertanto $f^*([(r_i^1)_i], \ldots, [(r_i^n)_i]) \downarrow \Leftrightarrow \left\{i \in \mathbb{N} \mid f(r_i^1, \ldots, r_i^n) \downarrow\right\} \in U$). Tuttavia, in generale non distingueremo notazionalmente tra $r, f, P$ e $r^*, f^*, P^*$, rispettivamente. Quindi ad esempio identificheremo $r \in \mathbb{R}$ con $[(r)_i] \in \mathbb{R}^*$. L'unica eccezione concerne i quantificatori (limitati), dove distingueremo tra l'insieme $A$ e l'insieme $A^* = \{[(r_i)_i] \mid \{i \mid r_i \in A\} \in U\}$.

Data una $\mathscr{L}_{\mathbb{R}}$-formula $\varphi$, la sua $*$-trasposta $\varphi^*$ sarà ottenuta sostituendo ciascun quantificatore $(Qx \in A)$ con il quantificatore $(Qx \in A^*)$.

Gli elementi di $\mathbb{R}^*$ sono detti *iperreali*.

Abbiamo che $[(r_i)_i] \leq [(r_i')_i] \Leftrightarrow \left\{i \in \mathbb{N} \mid r_i \leq r_i'\right\} \in U$ come anche $[(r_i)_i] < [(r_i')_i] \Leftrightarrow \left\{i \in \mathbb{N} \mid r_i < r_i'\right\} \in U$, ed è facile vedere che tali relazioni definiscono degli ordini lineari su $\mathbb{R}^*$ sulla base delle proprietà degli ultrafiltri. Vediamo ad esempio il caso della linearità per $<$ (il resto è lasciato per esercizio). Sia $[(r_i)_i] \neq [(r'_i)_i]$. Se $\{i \mid r_i < r'_i\} \in U$, allora $[(r_i)_i] < [(r'_i)_i]$ e siamo a posto. Altrimenti $\{i \mid r_i \geq r'_i\} \in U$, ma anche $\{i \mid r_i \neq r'_i\} \in U$, da cui $\{i \mid r_i \geq r'_i\} \cap \{i \mid r_i \neq r'_i\} = \{i \mid r_i > r'_i\} \in U$, quindi $[(r_i)_i] > [(r'_i)_i]$.

Attenzione! $[(0)_n], [(\frac{1}{2n})_n], [(\frac{1}{n})_n]$ sono iperreali ben disinti anche se $\lim_{n\to\infty} 0 = \lim_{n\to\infty} \frac{1}{2n} = \lim_{n\to\infty} \frac{1}{n}$, e infatti abbiamo che $[(0)_n] < [(\frac{1}{2n})_n] < [(\frac{1}{n})_n]$. Ponendo $\varepsilon := [(\frac{1}{2n})_n], \varepsilon' := [(\frac{1}{n})_n]$ abbiamo che $\varepsilon, \varepsilon'$ sono due numeri positivi, eppure più piccoli di ogni numero reale standard positivo $r \in \mathbb{R}$, in quanto $\{n \mid \frac{1}{kn} < r\} \in U$ per qualsiasi $k \in \mathbb{N}^+$ giacché $U$ è un filtro non principale. Pertanto $\varepsilon, \varepsilon'$ sono due *infinitesimi* positivi di diversa grandezza (0 lo consideriamo un infinitesimo banale).

Diciamo che due iperreali $[(r_i)_i], [(r'_i)_i]$ sono *infinitamente vicini*, o *a distanza infinitesima*, in simboli $[(r_i)_i] \simeq [(r'_i)_i]$, se per ogni $n \in \mathbb{N}$ l'insieme $\{i \mid |r_i - r'_i| < \frac{1}{n}\} \in U$. È facile vedere che $\simeq$ è una relazione di equivalenza (esercizio).

Possiamo quindi definire in questo modo per un qualsiasi reale standard $r \in \mathbb{R}$ la sua *monade*[4] $\mathrm{mon}(r) := \{[(r_i)_i] \mid [(r_i)_i] \simeq r\}$.

Come abbiamo gli infinitesimi, altrettanto abbiamo i numeri *illimitati*, ovvero quelli i cui valori assoluti sono più grandi di ogni numero reale dato: si considerino ad esempio $[(n)_n], [(-n)_n]$.

Un iperreale che non è illimitato è *limitato*.

Il Teorema di Łos (in una versione adatta al nuovo linguaggio e alla sua semantica) ci assicura la validità del seguente fondamentale principio per ogni $\mathscr{L}_{\mathbb{R}}$-formula $\varphi$ chiusa:

$$\textbf{Principio di transfer:}\quad \mathbb{R} \models \varphi \Longleftrightarrow \mathbb{R}^* \models \varphi^*$$

Il principio di transfer è veramente utile in Analisi Non Standard perché consente la dimostrazione di molti risultati in modo più semplice che ragionando in termini di ultrafiltri.

**Teorema 10.1** (parte standard)**.** *Ogni iperreale limitato b è infinitamente vicino ad uno ed un solo reale standard, chiamato la "parte standard" di b, ed indicato come* $\mathrm{st}(b)$.

[4]Leibniz è stato, insieme a Newton, uno degli inventori (scopritori?) del calcolo infinitesimale e il grande metafisico della teoria delle monadi.

*Dimostrazione.* Esistenza: consideriamo l'insieme

$$B := \{ r \in \mathbb{R} \mid r < b \} \subseteq \mathbb{R}.$$

Poiché $b$ è limitato, allora esistono due reali standard $r, s \in \mathbb{R}$ tali che $r < b < s$. L'esistenza di $r$ dimostra che $B \neq \emptyset$; quanto ad $s$, questo è un suo maggiorante. Poiché $\mathbb{R}$ è completo, $B$ ha un supremo $c$ in $\mathbb{R}$. Mostriamo che $b \simeq c$. Sia dato infatti un generico $d \in \mathbb{R}^+$. Allora $c < c + d$ e quindi $c + d \notin B$ in quanto $c$ è maggiore o uguale di tutti gli elementi in $B$, da cui $b \leq c + d$ per definizione di $B$. Poiché $c - d < c$ e $c$ è supremo di $B$ e $B$ è chiuso verso il basso (per transitività di $<$ su $\mathbb{R}^*$), allora $c - d \in B$ da cui $c - d < b$. Poiché in $\mathbb{R}$ si ha

$$(\forall x \in \mathbb{R})\,(c - d < x \leq c + d \to |x - c| \leq d)$$

per il principio del transfer in $\mathbb{R}^*$ si ha

$$(\forall x \in \mathbb{R}^*)\,(c - d < x \leq c + d \to |x - c| \leq d).$$

Pertanto, poiché $c - d < b \leq c + d$, abbiamo $|b - c| \leq d$. Ma $d \in \mathbb{R}^+$ è generico, da cui segue $b \simeq c$.

Unicità: $(c, c' \in \mathbb{R}$ & $c \simeq b \simeq c') \Rightarrow c \simeq c' \Rightarrow c = c'$ (perché $c, c' \in \mathbb{R}$). □

### *10.1.1 Continuità*

**Teorema 10.2.** *$f : \mathbb{R} \to \mathbb{R}$ è continua in $r \in \mathbb{R}$ sse $(x \simeq r \Rightarrow f(x) \simeq f(r))$.*

*Dimostrazione.* $\Rightarrow$) Sia fissato $\varepsilon \in \mathbb{R}^+$. Poiché $f$ è continua in $r$ esiste $\delta \in \mathbb{R}^+$ tale che

$$(\forall x \in \mathbb{R})(|x - r| < \delta \to |f(x) - f(r)| < \varepsilon)$$

è vera in $\mathbb{R}$. Per il principio del transfer, abbiamo in $\mathbb{R}^*$

$$(\forall x \in \mathbb{R}^*)(|x - r| < \delta \to |f(x) - f(r)| < \varepsilon).$$

Se $x \in \mathbb{R}^*$ è tale che $x \simeq r$ allora automaticamente $|x - r| < \delta$ e quindi $|f(x) - f(r)| < \varepsilon$. Poiché $\varepsilon \in \mathbb{R}^+$ è arbitrariamente piccolo, allora $f(x) \simeq f(r)$.

$\Leftarrow$) Sia $\varepsilon \in \mathbb{R}^+$. Prendiamo un qualsiasi $x \in \mathbb{R}^*$ tale che $|x - r| < d$ per un qualche infinitesimo $d$. Allora $x \simeq r$ e per ipotesi $f(x) \simeq f(r)$. Abbiamo pertanto in $\mathbb{R}^*$

$$(\exists y \in (\mathbb{R}^+)^*)(\forall x \in \mathbb{R}^*)(|x - r| < y \to |f(x) - f(r)| < \varepsilon).$$

Per il principio del transfer, in $\mathbb{R}$

$$(\exists y \in \mathbb{R}^+)(\forall x \in \mathbb{R})(|x - r| < y \to |f(x) - f(r)| < \varepsilon).$$

C'è pertanto un $\delta \in \mathbb{R}^+$ tale che in $\mathbb{R}$

$$(\forall x \in \mathbb{R})(|x - r| < \delta \to |f(x) - f(r)| < \varepsilon).$$

□

**Osservazione 10.3.** *Il Teorema 10.2 ci fornisce una giustificazione per cancellare* $4,9dt$ *nel calcolo della velocità instantanea* $\frac{ds}{dt}(1)$ *di caduta del grave. Poniamo che* $9,8+4,9dt$ *sia la velocità instantanea* $\frac{ds}{dt}$ *non in 1 ma in* $1+dt$ *(intuitivamente, imputiamo il valore* $9,8+4,9dt$ *non alla velocità all'inizio dell'intervallo temporale infinitesimo, quanto piuttosto alla velocità acquisita al termine di tale intervallo, il che pare più opportuno perché* $s(1+dt)-s(1)$ *misura lo spazio percorso dall'istante 1 fino al raggiugimento dell'istante* $1+dt$*, e quindi pare più naturale riferire lo spazio percorso* $ds$ *nell'intervallo temporale* $dt$ *a questo secondo istante). Ci aspettiamo ancora che il valore di* $\frac{ds}{dt}(1)$ *sia un numero reale, perché l'originaria funzione* $\frac{ds}{dt}$ *da cui siamo partiti è definita solo per numeri reali e quindi ha valori reali per argomenti reali. Giacché* $1 \simeq 1+dt$*, supponendo che* $\frac{ds}{dt}$ *sia continua, come ci aspettiamo, otteniamo che* $\frac{ds}{dt}(1) \simeq \frac{ds}{dt}(1+dt) = 9,8+4,9dt$*. Pertanto* $\frac{ds}{dt}(1)$ *dev'essere l'unico reale infinitamente vicino a* $9,8+4,9dt$*, e cioè* 9.8.

**Corollario 10.4.** *$f : D \to \mathbb{R}$ è continua per $D \subseteq \mathbb{R}$ sse per ogni $r \in D$ si ha che $x \simeq r \Rightarrow f(x) \simeq f(r)$.*

*Dimostrazione.* Esercizio. □

**Teorema 10.5.** *$f : D \to \mathbb{R}$ è uniformente continua per $D \subseteq \mathbb{R}$ sse per ogni $x, y \in D^*$ si ha che $x \simeq y \Rightarrow f(x) \simeq f(y)$.*

*Dimostrazione.* Esercizio. □

### *10.1.2 Successioni*

**Teorema 10.6.** $(s_n)_n \in \mathbb{R}^{\mathbb{N}}$ *converge a* $\ell \in \mathbb{R}$ *sse* $s_N \simeq \ell$ *per ogni* $N \in \mathbb{N}^*$ *illimitato*[5].

*Dimostrazione.* $\Rightarrow$) Sia $\lim_{n\to\infty} s_n = \ell$. Allora per ogni reale $\varepsilon \in \mathbb{R}^+$ esiste un $n \in \mathbb{N}$ tale che in $\mathbb{R}$

$$(\forall x \in \mathbb{N})(x > n \to |s_x - \ell| < \varepsilon).$$

Per transfer, in $\mathbb{R}^*$

$$(\forall x \in \mathbb{N}^*)(x > n \to |s_x - \ell| < \varepsilon).$$

Pertanto se $N \in \mathbb{N}^*$ è illimitato, allora $N > n$, da cui $|s_N - \ell| < \varepsilon$. Poiché $\varepsilon \in \mathbb{R}^+$ è arbitrariamente piccolo, si ha $s_N \simeq \ell$.

$\Leftarrow$) Sia $N \in \mathbb{N}^*$ illimitato. Allora per ipotesi, $s_N \simeq \ell$, e lo stesso vale per ogni altro $M \in \mathbb{N}^*$ illimitato, in particolare per ogni $M > N$. Dunque in $\mathbb{R}^*$ per ogni reale $\varepsilon \in \mathbb{R}^+$

$$(\forall x \in \mathbb{N}^*)(x > N \to |s_x - \ell| < \varepsilon),$$

$$(\exists y \in \mathbb{N}^*)(\forall x \in \mathbb{N}^*)(x > y \to |s_x - \ell| < \varepsilon).$$

Per transfer, in $\mathbb{R}$ si ha

$$(\exists y \in \mathbb{N})(\forall x \in \mathbb{N})(x > y \to |s_x - \ell| < \varepsilon).$$

Pertanto $\lim_{n\to\infty} s_n = \ell$. □

**Teorema 10.7.** $(s_n)_n \in \mathbb{R}^{\mathbb{N}}$ *è di Cauchy sse* $s_N \simeq s_M$ *per tutti gli* $N, M \in \mathbb{N}^*$ *illimitati.*

*Dimostrazione.* Esercizio. Si ricordi che $(s_n)_n$ è una successione di Cauchy se comunque si prenda un $\varepsilon > 0$ esiste un $n \in \mathbb{N}$ tale che per ogni $m, k \geq n$ si ha $|s_m - s_k| < \varepsilon$. □

I precedenti teoremi riguardavano il rapporto tra $\mathbb{R}$ ed $\mathbb{R}^*$ e la traduzione di concetti di $\mathbb{R}$ nelle loro controparti di $\mathbb{R}^*$. Tuttavia uno degli obiettivi fondamentali dell'analisi non standard è quello di esibire dimostrazioni (più semplici?) di risultati standard usando metrodi non standard.

[5] Le successioni sono estendibili ad $\mathbb{N}^*$ in quanto funzioni con dominio $\mathbb{N}$, ovvero la successione $(s_n)_n$ corrisponde alla funzione $s : \mathbb{N} \to \mathbb{R}$, per $s_n = s(n)$.

**Teorema 10.8.** *$(s_n)_n \in \mathbb{R}^{\mathbb{N}}$ è di Cauchy sse $(s_n)_n$ converge (a un punto) in $\mathbb{R}$ (ovvero $\mathbb{R}$ è uno spazio metrico completo).*

*Dimostrazione.* $\Rightarrow$) Sia $(s_n)_n$ successione di Cauchy in $\mathbb{R}$. Allora $(s_n)_n$ è limitata, ovvero esiste un qualche $b \in \mathbb{R}$ tale che

$$(\forall x \in \mathbb{N})\ s_x \leq b.$$

Per transfer,

$$(\forall x \in \mathbb{N}^*)\ s_x \leq b$$

dunque $s_N$ è limitato (da $b$) anche per $N$ illimitato, e quindi per Teorema 10.1 $\mathrm{st}(s_N)$ esiste ed è univocamente determinata. Per il Teorema 10.7 $s_N \simeq s_M$ per $M$ illimitato, e quindi $s_M \simeq \mathrm{st}(s_N)$, in quanto $\simeq$ è relazione di equivalenza. Per il Teorema 10.6, $(s_n)_n$ converge (a $st(s_N)$).

$\Leftarrow$) Sia $\lim_{n\to\infty} s_n = \ell$. Per il Teorema 10.6, $s_N \simeq \ell \simeq s_M$ per qualsiasi $N, M$ illimitati. Poiché $\simeq$ è relazione di equivalenza, allora $s_N \simeq s_M$. Da questo segue per Teorema 10.7 che $(s_n)_n$ è successione di Cauchy. □

## 10.2 Topologia non standard

**Teorema 10.9.**

1. *$O \subseteq \mathbb{R}$ è aperto sse per ogni $r \in O$ si ha che $x \simeq r \Rightarrow x \in O^*$, ovvero $\mathrm{mon}(r) \subseteq O^*$.*

2. *$A \subseteq \mathbb{R}$ è chiuso sse per ogni $r \in \mathbb{R}$, se $r \simeq x$ per qualche $x \in A^*$, allora $r \in A$, ovvero se $\mathrm{mon}(r) \cap A^* \neq \emptyset \Rightarrow r \in A$.*

*Dimostrazione.*

1. Esercizio.

2. $\Rightarrow$) Sia $A$ chiuso e sia $r \in \mathbb{R}$. Sia inoltre $r \simeq x \in A^*$. Allora in $\mathbb{R}^*$ per ogni standard $\varepsilon \in \mathbb{R}^+$ si dà

$$(\exists x \in A^*)\ |x - r| < \varepsilon.$$

Per transfer, in $\mathbb{R}$, per ogni $\varepsilon \in \mathbb{R}^+$ si ha

$$(\exists x \in A)\ |x - r| < \varepsilon$$

Pertanto $r$ è un punto di chiusura di $A$. Essendo $A$ chiuso, $r \in A$.

$\Leftarrow$) Sia $r \in \mathbb{R}$ un punto di chiusura di $A$. Allora in $\mathbb{R}$

$$(\forall y \in \mathbb{R}^+)(\exists x \in A)\, |x - r| < y.$$

Per transfer, in $\mathbb{R}^*$ si dà

$$(\forall y \in (\mathbb{R}^+)^*)(\exists x \in A^*)\, |x - r| < y.$$

Questo vale in particolare per gli $y$ infinitesimi. Dunque esiste un $x \in A^*$ tale che $r \simeq x$. Allora per ipotesi $r \in A$. Dunque $A$ è chiuso.

$\square$

**Proposizione 10.10.** *Per ogni* $r \in \mathbb{R}$, $\mathrm{mon}(r) = \bigcap \{ O^* \mid r \in O$ *e* $O$ *aperto*$\}$

*Dimostrazione.* $\mathrm{mon}(r) \subseteq \bigcap \{ O^* \mid r \in O$ e $O$ aperto$\}$ per Teorema 10.9.1. D'altronde, se $x \notin \mathrm{mon}(r)$ allora $x \not\simeq r$, da cui esiste un $\varepsilon \in \mathbb{R}^+$ tale che $|x - r| > \varepsilon$. Consideriamo ora $O := (r - \varepsilon, r + \varepsilon) \subseteq \mathbb{R}$. Allora in $\mathbb{R}$

$$(\forall y \in O)\, |y - r| < \varepsilon$$

da cui, per transfer, in $\mathbb{R}^*$

$$(\forall y \in O^*)\, |y - r| < \varepsilon.$$

Pertanto, sebbene $r \in O$, si ha che $x \notin O^*$. $\square$

**Teorema 10.11** (Criterio di Compattezza di Robinson)**.** *$K \subseteq \mathbb{R}$ è compatto sse per ogni $x \in K^*$ esiste $r \in K$ tale che $x \simeq r$, ovvero* $\mathrm{st}(x) \downarrow \in K$.

*Dimostrazione.* $\Rightarrow$) Fallisca il Criterio di Compattezza di Robinson per $K$. Allora esiste un iperreale $k \in K^*$ tale che per ogni $r \in K$ si dà $k \not\simeq r$. Per ogni tal $r$ consideriamo pertanto un reale standard $\varepsilon_r \in \mathbb{R}^+$ tale che $|k - r| \geq \varepsilon_r$. Allora $\{(r - \varepsilon_r, r + \varepsilon_r) \mid r \in K\}$ è ovviamente un ricoprimento aperto (in $\mathbb{R}$) di $K$, eppure non ha sottoricoprimenti finiti. Supponiamo infatti ad esempio che sia $K \subseteq \bigcup_{0 \leq i \leq n} (r_i - \varepsilon_{r_i}, r_i + \varepsilon_{r_i})$. Si osservi che in generale se $A \subseteq B$ allora $A^* \subseteq B^*$ e $(A \cup B)^* = A^* \cup B^*$ (mentre in generale $(\bigcup_{n \in \mathbb{N}} A_n)^* \neq \bigcup_{n \in \mathbb{N}} A_n^*$ (esercizio)). Pertanto $K^* \subseteq \bigcup_{0 \leq i \leq n} (r_i - \varepsilon_{r_i}, r_i + \varepsilon_{r_i})^*$. Perciò esiste un $i$, $0 \leq i \leq n$, tale che

$k \in (r_i - \varepsilon_{r_i}, r_i + \varepsilon_{r_i})^*$. Sia $B := (r_i - \varepsilon_{r_i}, r_i + \varepsilon_{r_i})$. Si osservi che vale in $\mathbb{R}$ che

$$(\forall x \in B)\, |x - r_i| < \varepsilon_{r_i}$$

da cui per transfer

$$(\forall x \in B^*)\, |x - r_i| < \varepsilon_{r_i},$$

e quindi $|k - r_i| < \varepsilon_{r_i}$, ma questo contraddice la definizione di $\varepsilon_{r_i}$.

$\Leftarrow$) Supponiamo che $K$ non sia compatto. Allora esiste una copertura $\mathscr{C}$ di insiemi aperti di $\mathbb{R}$ che non contiene nessuna copertura finita. Possiamo osservare che esiste una nuova copertura $\mathscr{C}'$ di $K$ costituita solo di intervalli aperti con estremi razionali e tale che ciascuno di questi intervalli razionali è contenuto in un qualche insieme aperto $O$ di $\mathscr{C}$. Sia $\mathscr{C}' := \{(q_n, q'_n) \mid n \in \mathbb{N}\}$ per opportune successioni $(q_n)_{n\in\mathbb{N}}, (q'_n)_{n\in\mathbb{N}} \in \mathbb{Q}^n$. Anche $\mathscr{C}'$ non contiene nessun ricoprimento aperto finito di $K$, perché altrimenti, fosse questo $\left\{(q_{i_j}, q'_{i_j}) \mid 1 \le j \le \ell\right\}$ avremmo che $\{O_j \mid 1 \le j \le \ell\}$ sarebbe un sottoricoprimento finito di $\mathscr{C}$ di $K$, dove $O_j$ è un aperto in $\mathscr{C}$ che contiene $(q_{i_j}, q'_{i_j})$. Quindi, per ogni $k \in \mathbb{N}$, $K \not\subseteq (q_0, q'_0) \cup \ldots \cup (q_k, q'_k)$. Pertanto in $\mathbb{R}$

$$(\forall y \in \mathbb{N})(\exists x \in K)(\forall z \in \mathbb{N})\,(z \le y \to \neg\, q_z < x < q'_z).$$

Per transfer, in $\mathbb{R}^*$

$$(\forall y \in \mathbb{N}^*)(\exists x \in K^*)(\forall z \in \mathbb{N}^*)\,(z \le y \to \neg\, q_z < x < q'_z).$$

Pertanto anche per $N \in \mathbb{N}^*$ illimitato esiste un $x \in K^*$ tale che $q_n < x < q'_n$ è falso per ogni $n \le N$, in particolare per tutti gli $n \in \mathbb{N}$. Ma tale $x$ non può essere infinitamente vicino a nessun $r \in K$, altrimenti, poiché $r \in (q_n, q'_n)$ per qualche $n \in \mathbb{N}$, allora avremmo $q_n < x < q'_n$ per Teorema 10.9.1. □

**Teorema 10.12.** *Sia $K \subseteq \mathbb{R}$ compatto. Se $f : K \to \mathbb{R}$ è continua allora è uniformemente continua.*

*Dimostrazione.* Per il Teorema 10.5 è sufficiente dimostrare che se $x \simeq y$ per $x, y \in K^*$ allora $f(x) \simeq f(y)$. Siano quindi $x, y \in K^*$ tali che $x \simeq y$. Essendo $K$ compatto, per il Criterio di Compattezza di Robinson Teorema 10.11 essi sono infinitamente vicini a qualche reale standard in $K$, e, per transitività di $\simeq$, ad uno stesso reale standard $r \in K$. Poiché $f$ è continua su $K$, è in particolare continua in $r$, per il Teorema 10.4 $f(x) \simeq f(r) \simeq f(y)$, da cui, per transitività di $\simeq$, $f(x) \simeq f(y)$. $\square$

**Teorema 10.13** (Heine-Borel)**.** *$K \subseteq \mathbb{R}$ è compatto sse $K$ è chiuso e limitato.*

*Dimostrazione.* $\Rightarrow$) Sia $K$ compatto. Vediamo innanzitutto che $K$ è limitato. Se così non fosse, varrebbe

$$(\forall x \in \mathbb{R}^+)(\exists y \in K)\ x < |y|.$$

Per transfer

$$(\forall x \in (\mathbb{R}^+)^*)(\exists y \in K^*)\ x < |y|.$$

Prendendo $x$ illimitato, si vede che ci sarà un $y \in K^*$ illimitato. Ma questo contraddice il Criterio di Compattezza di Robinson (esercizio).

Vediamo ora che $K$ è chiuso. Sulla base del Teorema 10.9.2 è sufficiente mostrare che per ogni $r \in \mathbb{R}$, se $r \simeq x$ per qualche $x \in K^*$ allora $r \in K$. Sia quindi $r \simeq x \in K^*$. Per il Criterio di Compattezza di Robinson, $x \simeq r'$ per qualche standard $r' \in K$. Ma per il Teorema 10.1, avremo che $r' = \mathrm{st}(x) = r$.

$\Leftarrow$) Essendo $K$ limitato, abbiamo in $\mathbb{R}$

$$(\forall x \in K)\ |x| \le b$$

per un opportuno $b \in \mathbb{R}$. Per transfer in $\mathbb{R}^*$

$$(\forall x \in K^*)\ |x| \le b.$$

Pertanto qualsiasi $x$ in $K^*$ è limitato, e quindi per il Teorema 10.1 $\mathrm{st}(x) \downarrow= r$ per un certo $r \in \mathbb{R}$, da cui $x \simeq r \in \mathbb{R}$. Essendo $K$ chiuso, per il Teorema 10.9.2 $r \in K$. Perciò $x \simeq r \in K$ e per il Criterio di Compattezza di Robinson $K$ è compatto. $\square$

**Teorema 10.14.** *Sia $K$ compatto e $f : K \to \mathbb{R}$ continua. Allora $f(K)$ è compatto.*

*Dimostrazione.* Sia $L := f(K)$. In $\mathbb{R}$ abbiamo

$$(\forall y \in L)(\exists x \in K)\ y = f(x).$$

Per transfer

$$(\forall y \in L^*)(\exists x \in K^*)\ y = f(x).$$

Sia pertanto $y \in L^*$ e $y = f(x)$ per $x \in K^*$. Essendo $K$ compatto, per il Criterio di Compattezza di Robinson $x \simeq r$ per qualche $r \in K$. Essendo $f$ continua in $K$, per Teorema 10.4 $y = f(x) \simeq f(r)$. Ma $f(r) \in f(K) = L \subseteq \mathbb{R}$. Per il Criterio di Compattezza di Robinson concludiamo che $f(K)$ è compatto. $\square$

# Bibliografia

Chang, C. C., & Keisler, H. J. (1973). *Model Theory.* Amsterdam: Elsevier.

Davis, M., & Hersh, R. (1972. Ristampa in “Logica”, Le Scienze Quaderni 60, a cura di C. Mangione, 1991). L’analisi non standard. *Le Scienze, 40.*

Goldblatt, R. (1998). *Lectures on the Hyperreals.* New York: Springer.

Marcja, A., & Toffalori, C. (2003). *A Guide to Classical and Modern Model Theory.* Dordrecht: Kluwer.

Marker, D. (2002). *Model Theory: an Introduction.* New York: Springer.

Mendelson, E. (1972). *Introduzione alla Logica Matematica.* Torino: Bollati Boringhieri.

Poizat, B. (2012). *A Course in Model Theory: an Introduction to Contemporary Mathematical Logic.* New York: Springer.