ALGEBRA I: IL LEMMA DI ZORN ED IL SUO UTILIZZO

1. Processi e costruzioni infinite

Molte volte, in matematica, c'è la necessità di ripetere una data costruzione infinite volte. In tale situazione è spesso necessario compiere delle scelte arbitrarie, anch'esse in quantità infinita. La liceità dell'atto di compiere un'infinità di scelte arbitrarie è un argomento dibattuto: dal punto di vista puramente logico è stato mostrato che supporre di poterlo fare non porta a contraddizioni — in altre parole, l'*assioma della scelta*, che garantisce la possibilità di compiere infinite scelte, è indipendente dagli altri assiomi generalmente usati in matematica.

L'assioma della scelta, insieme alle sue molteplici riformulazioni equivalenti, permette di mostrare molte proprietà interessanti in molte strutture algebriche; consente tuttavia di esibire anche comportamenti profondamente antiintuitivi: attraverso l'assioma della scelta si costruiscono[1] sottoinsiemi non misurabili di $\mathbb{R}$; si decompone[2] la palla unitaria di $\mathbb{R}^3$ in un numero finito di pezzi che possono essere risistemati, attraverso movimenti rigidi, per ricomporre due palle unitarie distinte.

Il lemma di Zorn è forse la riformulazione più duttile dell'assioma della scelta, anche se a primo impatto è un po' dura da digerire. Prima di enunciarlo, vi ricordo che una *relazione d'ordine* su un insieme X è una relazione riflessiva, antisimmetrica e transitiva. Se su X è data una relazione d'ordine $\leq$, l'insieme X, o meglio la coppia $(X, \leq)$, si dice allora *insieme parzialmente ordinato*.

Una relazione d'ordine su X può essere *totale* quando, per ogni scelta di $x, y \in X$, almeno una tra $x \leq y$ e $y \leq x$ è vera — chiaramente sono entrambe vere se e solo se $x = y$; tuttavia la maggior parte delle relazioni d'ordine che ci interesseranno non saranno totali. Può accadere invece che un sottoinsieme C di X sia totalmente ordinato rispetto a $\leq$: in tal caso, C è detto *catena*. E' importante comprendere come le catene non debbano essere necessariamente sottoinsiemi finiti, né tantomeno numerabili. Una catena è semplicemente un sottoinsieme nel quale tutti gli elementi sono confrontabili, e può essere grande quanto vogliamo.

Esempio: Sia $A = \{a, b, c, 1, 2\}$, e sia X il suo insieme delle parti. La relazione di inclusione $\subseteq$ è di ordine parziale, ma non totale, in X. Ad esempio, nessuno tra i due sottoinsiemi $\{a, b\}, \{b, 1, 2\}$ è incluso nell'altro, sebbene non siano uguali. Tuttavia X contiene sottoinsiemi (di X) totalmente ordinati. Ad esempio:

$$C = \{\emptyset, \{a\}, \{a, b, 1\}, \{a, b, 1, 2\}\}$$

è totalmente ordinato, poiché comunque presi due suoi elementi (che sono sottoinsiemi di A) uno dei due è contenuto nell'altro. C è una di quelle che abbiamo definito catene: X non è totalmente ordinato da $\subseteq$, ma $C \subset X$ sì.

Vi ricordo ancora che, in un insieme parzialmente ordinato $(X, \leq)$, si chiama *maggiorante* di $Y \subset X$ ogni elemento $m \in X$ tale che $y \leq m$ per ogni $y \in Y$. Ad esempio 2 è un maggiorante di $Y = (0, 1)$ in $X = (\mathbb{R}, \leq)$ — a dire il vero ogni $m \geq 1$ è un maggiorante di Y. Un elemento $x \in X$ è invece *massimale* in X se non ci sono in X elementi più grandi, cioè se $x \leq y \Rightarrow x = y$. Ogni insieme parzialmente ordinato non vuoto *finito* ammette elementi massimali: se così non fosse, sarebbe possibile costruire una catena infinita di elementi distinti ognuno $\leq$ del successivo. Siamo pronti ad enunciare il

Lemma di Zorn: Sia $(\mathcal{F}, \leq)$ un insieme parzialmente ordinato non vuoto nel quale ogni catena ha (almeno) un maggiorante. Allora $\mathcal{F}$ possiede elementi massimali.

Se credete che ogni insieme parzialmente ordinato debba contenere elementi massimali, pensate all'insieme $\mathcal{F}$ i cui elementi sono i sottoinsiemi finiti di $\mathbb{N}$, ordinato rispetto all'inclusione. Chiaramente nessun elemento di $\mathcal{F}$ è massimale, perché a ogni sottoinsieme finito di $\mathbb{N}$ posso aggiungere un elemento, ottenendo così un sottoinsieme più grande, ma ancora finito.

Questo insieme $\mathcal{F}$ non contiene elementi massimali, e non può quindi soddisfare le ipotesi del Lemma di Zorn: deve ammettere catene senza maggioranti. Ad esempio, se C è il sottoinsieme di $\mathcal{F}$ i cui elementi sono tutti i sottoinsiemi della forma $\{0, 1, \ldots, n\}$:

$$C = \{\{0\}, \{0, 1\}, \{0, 1, 2\}, \{0, 1, 2, 3\}, \ldots\},$$

allora C è chiaramente una catena che non ammette alcun maggiorante in $\mathcal{F}$. In effetti, un sottoinsieme di $\mathbb{N}$ che contenga tutti tali sottoinsiemi (che sono tutti finiti) dovrebbe essere $\mathbb{N}$ stesso, che non è un insieme finito, e quindi non è un elemento di $\mathcal{F}$.

Nonostante il nome del Lemma di Zorn, noi lo prenderemo come principio da non dimostrare, cioè come assioma. In effetti, come vedremo in seguito, può essere dimostrato a partire dall'Assioma della scelta, ma l'Assioma della scelta stesso segue a partire dal Lemma di Zorn: in altre parole, l'uno vale l'altro!

[1] Cercate "Insieme di Vitali" in rete.

[2] Keyword: paradosso di Banack-Tarski.

2. Il Lemma di Zorn e l'Assioma della scelta

Ho fatto un gran parlare, finora, dell'Assioma della scelta, ma non ho ancora detto che cosa sia:

Assioma della scelta: Sia I un insieme (di indici), ed $\mathcal{X} = \{X_i, i \in I\}$ una famiglia di insiemi (indicizzati da I); indichiamo inoltre con X l'unione di tutti gli X_i. Allora esiste una *funzione di scelta*, cioè un'applicazione $f : I \to X$ tale che $f(i) \in X_i$ per ogni $i \in I$.

Per i pignoli, avrei potuto utilizzare come insieme di indici $\mathcal{X}$ stesso, ed indicare l'unione di tutti gli elementi di $\mathcal{X}$ con $\bigcup \mathcal{X}$. Però garantire l'esistenza di $f : \mathcal{X} \to \bigcup \mathcal{X}$ tale che $f(x) \in x$ per ogni $x \in \mathcal{X}$ mi sembrava davvero troppo criptico! Quella data sopra non è l'unica formulazione possibile dell'assioma della scelta, ma una delle più naturali — e in ogni caso, sono tutte equivalenti.

Perché l'Assioma della scelta dovrebbe essere intuitivamente valido? Dal mio punto di vista[3], questo è chiaro: devo scegliere un elemento da uno degli X_i, un altro elemento da un altro degli X_i, e così via. E' chiaro che se le scelte le devo fare io, non termino mai; ma è altrettanto chiaro che una scelta di un elemento da ogni insieme è possibile — almeno a me è chiaro e intuitivo: non so a voi!

Il Lemma di Zorn è lo strumento creato appositamente per trasformare le parole "e così via" in un argomento stringente. Descrivo la dimostrazione che segue con estrema attenzione ai dettagli, perché è il prototipo di ogni utilizzo del Lemma di Zorn. Le dimostrazioni che fanno uso del Lemma di Zorn diverranno sempre più asciutte, man mano che diventeremo familiari con tale strumento.

Dimostrazione dell'Assioma della scelta a partire dal Lemma di Zorn: Definiamo un insieme F come segue

$$\mathcal{F} = \{(J, f) \mid \quad J \subset I, \quad f : J \to X \text{ è tale che } f(i) \in X_i \text{ per ogni } i \in J\}.$$

In altre parole, $\mathcal{F}$ è l'insieme delle funzioni di scelta *parziali*, cioè di quelle funzioni che scelgono un elemento da ciascun X_i non per tutti gli $i \in I$, ma solo per quegli i che appartengono ad un sottoinsieme $J \subset I$.

Innanzitutto, l'insieme $\mathcal{F}$ è non vuoto: sia perché esiste una funzione di scelta parziale definita su $J = \emptyset$, sia perché compiere una quantità finita di scelte non crea problemi a nessuno, e quindi esistono anche funzioni di scelta parziali definite su sottoinsiemi finiti di I. Possiamo inoltre definire una relazione di ordine parziale su $\mathcal{F}$ come segue: $(J, f) \leq (J', f')$ se e solo se $J \subset J'$ e la restrizione di f' a J coincide con f. In altri termini $(J, f) \leq (J', f')$ se f' è sicuramente definita su tutti gli indici sui quali è definita anche la f (ma possibilmente anche su altri indici), e su tali indici sceglie gli stessi elementi che sceglie f: in parole povere $(J, f) \leq (J', f')$ se f' *estende* f.

Ora, se un elemento (J, f) in $\mathcal{F}$ non è definito su tutto I, cioè $J \neq I$, è facile estenderlo ad un insieme un po' più grande: si sceglie $i \notin J$, e si sceglie $f(i) \in X_i$. Queste sono solo due scelte da fare, e non rappresentano una difficoltà psicologica insormontabile. Ci siamo quindi convinti che (J, f) non possa essere massimale in $\mathcal{F}$, a meno che $J = I$. Ma se $(I, f) \in \mathcal{F}$, allora $f : I \to X$ è una funzione di scelta! Quindi per mostrare l'esistenza di una funzione di scelta è sufficiente mostrare l'esistenza di elementi massimali in $\mathcal{F}$. E' qui che entra in gioco il Lemma di Zorn.

Il Lemma di Zorn garantisce l'esistenza di elementi massimali in $\mathcal{F}$ non appena siamo in grado di mostrare che ogni catena in $\mathcal{F}$ ammette un maggiorante. Sia quindi C una catena in $\mathcal{F}$: gli elementi di C sono coppie (J, f) tutte confrontabili tra loro; le funzioni di scelta parziali corrispondenti si estendono l'una con l'altra. Ogni maggiorante di C deve essere una coppia $(\overline{J}, \overline{f})$ con la proprietà che $J \subset \overline{J}$ per ogni $(J, f) \in C$ e tale che $\overline{f}$ estende tutte le f contemporaneamente. Ma costruire un tale maggiorante è facile!

Si prende come $\overline{J}$ l'unione di tutti i J degli elementi di C, e si definisce $f : \overline{J} \to X$ come $\overline{f}(j) = f(j)$ se $(f, J) \in C$ e $j \in J$. Questa definizione non dipende dalla scelta di $(f, J) \in C$ perché gli elementi di C si estendono l'uno con l'altro. Inoltre $\overline{J}$ è l'unione di tutti i J degli elementi di C, e quindi se $j \in \overline{J}$, allora j appartiene ad almeno uno dei sottoinsiemi J.

Abbiamo mostrato che ogni catena in $\mathcal{F}$ ammette un maggiorante; grazie al Lemma di Zorn, $\mathcal{F}$ possiede elementi massimali, cioè funzioni di scelta per la famiglia $\mathcal{X} = \{X_i, i \in I\}$. □

L'utilizzo del Lemma di Zorn si fa sempre in questo modo: si inventa un insieme parzialmente ordinato i cui elementi massimali diano risposta positiva al nostro problema; quindi si costruisce un maggiorante per ogni catena. Nel caso dell'Assioma della scelta, le funzioni di scelta sono funzioni di scelta parziali massimali (va mostrato, e noi lo abbiamo mostrato), e l'esistenza di maggioranti delle catene si fa semplicemente considerando la funzione di scelta definita sull'unione dei domini delle funzioni di scelta parziali appartenenti alla catena.

Notate che la dimostrazione di sopra traduce perfettamente la dimostrazione data a lezione, che era: *scelgo un indice, e scelgo un elemento dall'insieme che indicizza, poi scelgo un altro indice, e scelgo un elemento dall'insieme che indicizza... Quando non posso più andare avanti, vuol dire che l'insieme di indici per i quali ho operato la scelta coincide con tutto I.*

Il Lemma di Zorn è quel che c'è dietro i puntini di sospensione: nel procedimento di scegliere ogni volta un nuovo indice, ed un elemento dall'insieme che indicizza, sto costruendo una catena di funzioni di scelta parziali. Come esseri umani, possiamo operare soltanto una famiglia finita, arbitrariamente grande, di scelte (quindi costruire una catena numerabile), ma allora il Lemma di Zorn ci garantisce l'esistenza di un maggiorante, cioè di una scelta fatta sull'insieme numerabile (grande almeno quanto quello) dato dall'unione di tutti gli indici che abbiamo considerato finora. Ma allora possiamo scegliere un altro indice fuori, ed un altro elemento nell'insieme che indicizza, e continuare la nostra catena oltre l'infinità numerabile di scelte fatta inizialmente. Anche in questo caso, il Lemma di Zorn

[3]Ma come vi ho detto, la percezione di che cosa sia *intuitivo* varia da persona a persona..

ci garantisce l'esistenza di una funzione di scelta definita sull'unione dei due insiemi numerabili, e di poter andare avanti.

La cosa stupefacente è che il Lemma di Zorn incorpora al suo interno la possibilità, procedendo di scelte numerabili in scelte numerabili, di raggiungere sottoinsiemi di cardinalità qualsivoglia elevata: l'importante è essere in grado, ad ogni passo, di costruire un maggiorante (cioè una estensione collettiva di tutte le funzioni di scelta compatibili fino a quel momento considerate) di qualsiasi catena, qualsiasi sia la sua cardinalità.

3. Forme equivalenti dell'assioma della scelta

3.1. **Gergo.** Ricapitoliamo ora per comodità tutte le definizioni già date. Una relazione $\leq$ sull'insieme X si dice *ordinamento parziale* se soddisfa:

- $x \leq x$ per ogni $x \in X$ *Riflessività*
- $x \leq y, y \leq x \implies x \leq z$ *Transitività*
- $x \leq y, y \leq x \implies x = y$ *Antisimmetria*

Ad esempio, l'inclusione $\subseteq$ è un'ordinamento parziale sull'insieme $X = P(\Omega)$ delle parti di un insieme Ω dato. E' importante notare come, dati $x, x' \in X$, non si richiede che $x \leq x'$ oppure $x' \leq x$: se questo accade, x, x' si dicono *confrontabili*. Un ordinamento parziale per il quale ogni coppia di elementi sia confrontabile si dice *ordinamento totale*. Se $x \leq x'$, si scrive anche $x' \geq x$; se $x \leq x'$ e $x \neq x'$, si scrive anche $x < x'$ o equivalentemente $x' > x$. Attenzione! Un ordinamento totale **è** un ordinamento parziale. Un insieme dotato di un ordinamento parziale/totale si dice, ovviamente, *insieme parzialmente/totalmente ordinato*.

Esempio 3.1. Le naturali relazioni d'ordine su $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ sono tutte ordinamenti totali.

Esempio 3.2. Se X possiede al più un elemento, ogni ordinamento parziale su X è totale.

Se $U \subset X$, un *maggiorante* (risp. *minorante*) di U è ogni elemento $m \in X$ tale che $m \geq u$ (risp. $m \leq u$) per ogni $u \in U$: in generale, $U \subset X$ può avere più di un maggiorante in X, ma può anche non averne alcuno.

Un elemento $m \in X$ si dice *massimale* (risp. *minimale*) se $m \leq m'$ (risp. $m \geq m'$) implica $m = m'$. Si dice *massimo* (risp. *minimo*) se $m \geq x$ (risp. $m \leq x$) per ogni $x \in X$. Un massimo (minimo) di X è sempre unico; ogni massimo (minimo) di X è sempre massimale (minimale) in X, ma il contrario non è necessariamente vero; inoltre, un insieme parzialmente ordinato può avere più di un elemento massimale (minimale) o può anche non averne nessuno.

Esempio 3.3. L'unico elemento minimale di $(\mathbb{N}, \leq)$ è 0, che è anche il suo elemento minimo. Al contrario, $(\mathbb{N}, \leq)$ non ha elementi massimali (e quindi neanche massimi).

Gli insiemi $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$, rispetto alla naturale relazione d'ordine, non hanno elementi massimali né minimali.

Se $(X, \leq)$ è un insieme parzialmente ordinato e $Y \subset X$, allora la restrizione a Y di $\leq$ è ancora una relazione d'ordine parziale: questo ci permette di interpretare, cosa che faremo sempre, i sottoinsiemi di un insieme parzialmente ordinato come insiemi parzialmente ordinati.

Può capitare che un sottoinsieme Y di $(X, \leq)$ sia totalmente ordinato rispetto all'ordinamento ereditato da X, anche se X non lo è necessariamente. In tal caso, diremo che Y è una *catena* in X.

3.2. **Insiemi bene ordinati.** Ogni sottoinsieme non vuoto di $\mathbb{N}$, rispetto alla naturale relazione d'ordine, possiede un minimo elemento: questo fatto è noto come *principio di buon ordinamento* ed è essenzialmente equivalente al principio di induzione.

Definizione 3.4. Un ordinamento parziale $\leq$ su X si dice *buon ordinamento* se ogni sottoinsieme non vuoto di X possiede minimo.

Un insieme dotato di un buon ordinamento si dice *bene ordinato*. Se x, x' sono elementi distinti di un insieme bene ordinato $(X, \leq)$, allora il minimo del sottoinsieme non vuoto $\{x, x'\} \subset X$ è minore o uguale dell'altro elemento. Pertanto ogni buon ordinamento è un ordinamento totale. Il viceversa è però falso.

Esempio 3.5. Il sottoinsieme $(0, 1) \subset \mathbb{R}$ non ha minimo rispetto alla naturale relazione d'ordine. Pertanto $(\mathbb{R}, \leq)$ è un insieme totalmente ordinato ma non bene ordinato.

Esempio 3.6. Ogni ordinamento totale su un insieme finito è un buon ordinamento. In particolare, ogni ordinamento parziale su un insieme con al più un elemento è un buon ordinamento.

Proposizione 3.7. *Sia $(X, \leq)$ è un insieme bene ordinato, $y \notin X$. Allora l'ordinamento su $X' = X \cup \{y\}$ che estende quello di X ponendo $x \leq y$ per ogni $x \in X'$ è ancora un buon ordinamento.*

Dimostrazione. La verifica che si tratti di un ordinamento parziale è immediata, ed è quindi sufficiente mostrare che ogni sottoinsieme non vuoto di X' ha elemento minimo.

Se $\emptyset \neq U \subset X'$ non contiene y, questo segue dal fatto che X è bene ordinato. Se invece lo contiene propriamente, il minimo di $U \setminus \{y\}$ è anche minimo di U. Se infine $U = \{y\}$, allora y è chiaramente il suo minimo. □

Se $(X, \leq)$ è un insieme bene ordinato e $Y \subset X$, allora anche Y è bene ordinato[4] dalla restrizione a Y di $\leq$. Il fatto che un ordinamento parziale su X induca buoni ordinamenti su alcuni sottoinsiemi dice però poco sul fatto che $\leq$ sia un buon ordinamento anche su X, come mostrano i seguenti esempi.

[4]ed è quindi automaticamente una catena!

Esempio 3.8.
- L'unione di sottoinsiemi bene ordinati può non essere bene ordinata. In effetti, ogni $(X, \leq)$ è unione dei suoi sottoinsiemi di cardinalità 1, che sono tutti bene ordinati, ma un ordinamento parziale è raramente un buon ordinamento.
- L'unione crescente di sottoinsiemi bene ordinati può non essere bene ordinata. In effetti, per ogni scelta di $k \in \mathbb{Z}$ il sottoinsieme $X_k = \{z \in \mathbb{Z} \mid z \geq k\}$ è bene ordinato rispetto all'ordinamento naturale di $\mathbb{Z}$; inoltre gli insiemi $X_k, k \in \mathbb{Z}$, sono contenuti l'uno nell'altro. Tuttavia la loro unione $(\mathbb{Z}, \leq)$ non è un insieme bene ordinato.

Esiste però un modo di ovviare a questo problema.

Definizione 3.9. Sia $(X, \leq)$ un insieme parzialmente ordinato. Il sottoinsieme $U \subset X$ si dice *segmento iniziale* di X quando $u \in U, x \leq u \implies x \in U$.

Esempio 3.10. Ciascun sottoinsieme $\{0, 1, \ldots, n\}$ è un segmento iniziale di $(\mathbb{N}, \leq)$. Nessuno dei sottoinsiemi X_k dell'esempio precedente è un segmento iniziale di $(\mathbb{Z}, \leq)$.

Quando $(X, \leq)$ è un insieme bene ordinato, i segmenti iniziali di X hanno una descrizione molto semplice.

Lemma 3.11. *Sia $(X, \leq)$ un insieme bene ordinato. Se $U \subsetneq X$ è un segmento iniziale di X, allora esiste $x_0 \in X$ tale che $U = X_{<x_0} := \{x \in X \mid x < x_0\}$.*

Dimostrazione. Se $U \subsetneq X$, sia $x_0 = \min X \setminus U$. Se $x < x_0$ allora $x \in U$ per minimalità di x_0. Viceversa, se $x \in U, x_0 < x$, allora $x_0 \in U$ in quanto U è un segmento iniziale; ma questo è un assurdo. □

Una rapida osservazione: se $U \neq V$ sono bene ordinati e U è un segmento iniziale di V, allora $u < v$ per ogni $u \in U, v \in V \setminus U$. In effetti, sia $v_0 = \min V \setminus U$. Se $u \in U$, allora $v_0 \leq u$ è impossibile, perché U è un segmento iniziale di V e questo obbligherebbe $v_0 \in U$. Pertanto, se $v \in V \setminus U$ abbiamo $u < v_0 \leq v$ e quindi $u < v$.

Proposizione 3.12. *Sia $(X, \leq)$ un insieme parzialmente ordinato e $\{U_i, i \in I\}$ sottoinsiemi bene ordinati con la proprietà che comunque presi $i, j \in I$, uno tra U_i e U_j è segmento iniziale dell'altro. Allora l'unione $U = \bigcup_{i \in I} U_i$ è ancora un sottoinsieme bene ordinato di X.*

Dimostrazione. Se $Y \subset U$ è un sottoinsieme non vuoto, esiste sicuramente $i \in I$ tale che $Y \cap U_i \neq \emptyset$. Indichiamo con y_0 il minimo di $Y \cap U_i$ che esiste in quanto U_i è bene ordinato. Sia $y \in Y$: se $y \in U_i$ allora $y \in Y \cap U_i$ e quindi $y_0 \leq y$; se invece $y \notin U_i$, scegliamo $j \in I$ in modo che $y \in U_j$. Allora $U_j \setminus U_i$ è non vuoto perché contiene y e quindi U_i è un segmento iniziale di U_j. Per l'osservazione appena fatta, $U_i \ni y_0 < y \in U_j \setminus U_i$. In conclusione, y_0 è il minimo di Y. □

3.3. **Enunciati equivalenti all'assioma della scelta.** Elenchiamo adesso alcuni enunciati dei quali vogliamo mostrare l'equivalenza.

Assioma della scelta: prima forma. Se X, Y sono insiemi e X è non vuoto, allora per ogni applicazione suriettiva $f : X \to Y$ esiste un'applicazione $g : Y \to X$ tale che $f \circ g = \mathrm{id}_Y$.

Assioma della scelta: seconda forma. Sia I un insieme e $\{X_i, i \in I\}$ una famiglia di insiemi non vuoti. Allora esiste un'applicazione $\varphi : I \to \bigcup_{i \in I} X_i$ tale che $\varphi(i) \in X_i$.

Principio di buon ordinamento. Ogni insieme X possiede un buon ordinamento.

Lemma di Zorn. Un insieme parzialmente ordinato non vuoto, ogni cui catena possiede un maggiorante[5], ha necessariamente elementi massimali.

Abbiamo già visto che il Lemma di Zorn permette di mostrare direttamente la seconda forma dell'assioma della scelta, ma procederemo ora in maniera diversa. Iniziamo dimostrando che dal Lemma di Zorn segue il Principio di buon ordinamento ed immediatamente anche la prima forma dell'assioma della scelta.

Proposizione 3.13. *Se ogni insieme parzialmente ordinato non vuoto induttivo ha elementi massimali, allora ogni insieme può essere bene ordinato.*

Dimostrazione. Vogliamo costruire, utilizzando il Lemma di Zorn, un buon ordinamento sull'insieme X, comunque sia dato. L'insieme

$$\mathcal{F} = \{(U, \leq_U) \mid U \subset X,\ \leq_U \text{ è un buon ordinamento su } U\}$$

è sicuramente non vuoto, poiché contiene la coppia $(\emptyset, \emptyset)$. Si vede facilmente che la relazione su $\mathcal{F}$ definita da

$$(U, \leq_U) \preceq (V, \leq_V) \iff U \text{ è segmento iniziale di } V, \text{ con la relazione d'ordine ereditata da } V$$

è un ordinamento parziale su $\mathcal{F}$.

Mostriamo che $\mathcal{F}$ è un insieme induttivo. Se $\{U_i, i \in I\}$ è una catena in $\mathcal{F}$, sia $U = \bigcup_{i \in I} U_i$. Se $u, u' \in U$, possiamo trovare $i \in I$ in modo che U_i li contenga entrambi. Definiamo allora $u \leq_U u'$ se $u \leq_{U_i} u'$: questa definizione non dipende da i in quanto gli ordinamenti dei sottoinsiemi $U_i, i \in I$, coincidono sugli elementi comuni. Pertanto

[5]Un insieme parzialmente ordinato che soddisfa questa ulteriore condizione è detto *induttivo*.

$(U, \leq_U)$ è un insieme parzialmente ordinato e possiamo usare la Proposizione 3.12 per mostrare che $\leq_U$ è un buon ordinamento su U.

Rimane da mostrare che ogni U_i è un segmento iniziale di U: se $x \in U, u \in U_i$ e $x \leq u$, allora x deve appartenere ad U_i. In effetti, se $x \notin U_i$, allora $x \in U_j$ per qualche $j \in I$ e poiché $U_j \setminus U_i$ è non vuoto (in quanto contiene x), allora U_i è segmento iniziale di U_j e gli elementi di U_i sono tutti inferiori a quelli di $U_j \setminus U_i$: $x \leq u$ sarebbe quindi impossibile. Questo mostra che $(U_i, \leq_{U_i}) \preceq (U, \leq_U)$ per ogni $i \in I$, e $(U, \leq_U)$ è quindi un maggiorante della catena data.

Essendo $\mathcal{F}$ un insieme parzialmente ordinato non vuoto e induttivo, deve possedere almeno un elemento massimale $(Y, \leq_Y)$. Tuttavia, se $Y \subsetneq X$, e $x \in X \setminus Y$, allora possiamo costruire $Y' = Y \cup \{x\}$ come nella Proposizione 3.7, e confutare la massimalità di $(Y, \leq_Y)$. Gli elementi massimali di $\mathcal{F}$ costituiscono quindi buoni ordinamenti su X. □

Corollario 3.14. *Supponiamo che ogni insieme possieda un buon ordinamento. Allora ogni suriezione $f : X \to Y$ ammette un'inversa destra.*

Dimostrazione. Se $\leq$ è un buon ordinamento su X, l'applicazione $g : Y \to X$ definita da $g(y) := \min f^{-1}(y)$ è un'inversa destra di f. □

Passare dalla prima forma alla seconda dell'assioma della scelta richiede solo di risolvere alcuni semplici aspetti tecnici.

Proposizione 3.15. *Supponiamo di sapere che ogni suriezione tra insiemi possiede un'inversa destra. Allora ogni famiglia $\{X_i, i \in I\}$ di insiemi non vuoti indicizzati dall'insieme I possiede una funzione di scelta $\varphi : I \to \bigcup_{i \in I} X_i$ tale che $\varphi(i) \in X_i$.*

Dimostrazione. Se $X = \bigcup_{i \in I} X_i$ indichiamo con $\pi_1 : X \times I \to X$, $\pi_2 : X \times I \to I$ le proiezioni sulla prima e sulla seconda coordinata. Poniamo $A = \{(x, i) \in X \times I \mid x \in X_i\}$.

La restrizione $\pi_2|_A : A \to I$ è suriettiva poiché ogni X_i è non vuoto. Se $g : I \to A$ è la sua inversa destra, la composizione $\varphi = \pi_1|_A \circ g : I \to X$ è la funzione di scelta cercata. □

Rimane da dimostrare che il Lemma di Zorn segue dalla seconda forma dell'assioma della scelta. Questo è il punto più delicato e richiederà un po' di attenzione.

3.4. **L'assioma della scelta implica il Lemma di Zorn.** In tutto ciò che segue, $\mathcal{F}$ è un insieme parzialmente ordinato non vuoto e induttivo **privo** di elementi massimali. Poiché per ogni $x \in \mathcal{F}$ possiamo trovare $x' \in \mathcal{F}$ tale che $x < x'$, ogni catena $C \subset \mathcal{F}$ non solo possiede almeno un maggiorante, ma anche un *maggiorante stretto*, un elemento m cioè tale che $c < m$ per ogni $c \in C$.

Se I è l'insieme di tutte le catene in $\mathcal{F}$, e $X_i, i \in I$, è l'insieme dei maggioranti stretti di i in $\mathcal{F}$, sia $f : I \to \bigcup_{i \in I} X_i$ una funzione di scelta: f sceglie, per ogni catena $i \in I$, un suo maggiorante stretto $f(i) \in \mathcal{F}$.

Definizione 3.16. Un sottoinsieme $A \subset \mathcal{F}$ si dice f-sottoinsieme se

- A è bene ordinato dalla relazione d'ordine di $\mathcal{F}$;
- per ogni $a \in A$ vale $f(A_{<a}) = a$.

Non è complicato esibire f-sottoinsiemi. Innanzitutto, il sottoinsieme vuoto è tecnicamente un f-sottoinsieme. Inoltre, se $x = f(\emptyset)$, allora $\{x\}$ è un f-sottoinsieme, come si verifica facilmente. Continuando, se $y = f(\{x\})$, allora $\{x, y\}$ è nuovamente un f-sottoinsieme.

Più in generale, se un f-sottoinsieme $A \subset \mathcal{F}$ è chiaramente una catena di $\mathcal{F}$. Allora, detto $x = f(A)$, il sottoinsieme $A' = A \cup \{x\}$ è nuovamente un f-sottoinsieme: che sia bene ordinato segue dalla Proposizione 3.7, mentre la validità della seconda condizione è immediata.

Lemma 3.17. *Se $A \neq B$ sono f-sottoinsiemi di $\mathcal{F}$, allora uno dei due è segmento iniziale dell'altro.*

Dimostrazione. Innanzitutto, A e B sono entrambi bene ordinati. A meno di scambiare A con B, supponiamo che B non sia contenuto in A e poniamo $b_0 = \min B \setminus A$. Allora $B_{<b_0} \subset A$. Se quest'inclusione è un'uguaglianza, abbiamo finito.

Altrimenti, sia $a_0 = \min A \setminus B_{<b_0}$. Come prima, $A_{<a_0} \subset B_{<b_0}$: vogliamo mostrare che questa inclusione è ora effettivamente un'uguaglianza. Sia $b \in B_{<b_0}$. Se $b \geq a_0$, allora $a_0 \leq b < b_0$ mostrerebbe che $a_0 < b_0$ il che contraddice la definizione di a_0. Poiché l'ordinamento su A è totale, deve allora valere $b < a_0$, il che dimostra $B_{<b_0} \subset A_{<a_0}$, e quindi l'uguaglianza, dal momento che l'inclusione opposta era già nota.

Pertanto $A_{<a_0} = B_{<b_0}$. Ricordando che A, B sono entrambi f-sottoinsiemi, abbiamo allora

$$A \ni a_0 = f(A_{<a_0}) = f(B_{<b_0}) = b_0,$$

il che è assurdo, poiché b_0, per definizione, non appartiene ad A. □

Lemma 3.18. *L'unione di tutti gli f-sottoinsiemi di $\mathcal{F}$ è ancora un f-sottoinsieme.*

Dimostrazione. Poiché f-sottoinsiemi distinti sono uno segmento iniziale dell'altro, il fatto che l'unione di tutti gli f-sottoinsiemi di $\mathcal{F}$ sia bene ordinata segue dalla Proposizione 3.12.

Per quanto riguarda la seconda condizione, sia M l'unione di tutti gli f-sottoinsiemi di $\mathcal{F}$. Se $m \in M$, allora $m \in A$ per qualche f-sottoinsieme $A \subset \mathcal{F}$. Sia $x \in M$: se $x \notin A$, allora $x \in B$ per qualche altro f-sottoinsieme B. Poiché $B \setminus A$ è non vuoto, allora A è segmento iniziale di B e quindi $x > m$. Questo mostra che gli elementi di M

che sono minori di m sono tutti contenuti in A e più precisamente che $M_{<m} = A_{<m}$. Poiché A è un f-sottoinsieme, abbiamo allora

$$f(M_{<m}) = f(A_{<m}) = m,$$

e anche M è un f-sottoinsieme di $\mathcal{F}$. □

Proposizione 3.19. *Se ogni insieme di insiemi non vuoti ammette una funzione di scelta, allora ogni insieme parzialmente ordinato non vuoto e induttivo possiede elementi massimali.*

Dimostrazione. Supponiamo (per assurdo) che esista un insieme parzialmente ordinato non vuoto e induttivo $\mathcal{F}$ privo di elementi massimali. Se f è una funzione di scelta che associa ad ogni catena in $\mathcal{F}$ un suo maggiorante stretto, allora l'unione M degli f-sottoinsiemi di $\mathcal{F}$ è un f-sottoinsieme di $\mathcal{F}$ che contiene ogni altro f-sottoinsieme.

Tuttavia, come abbiamo già visto, $M \cup f(M)$ è ancora un f-sottoinsieme, il che contraddice la massimalità di M tra gli f-sottoinsiemi di $\mathcal{F}$. □

4. Esempi di utilizzo del Lemma di Zorn

4.1. **Esistenza di ideali massimali.** Se A è un anello con 1, mostriamo che esistono necessariamente ideali sinistri massimali in A. Consideriamo l'insieme $\mathcal{F} = \{I \subset A |\ I \text{ è un ideale sinistro proprio di } A\}$ e ordiniamolo per inclusione. Si vede subito che $\mathcal{F}$ è non vuoto, poiché $(0) \in \mathcal{F}$.

Lemma 4.1. *L'insieme $\mathcal{F}$ è induttivo.*

Dimostrazione. Sia $C = \{I_k, k \in K\}$ una catena in $\mathcal{F}$. Vogliamo mostrare che $I = \bigcup_{k\in K} I_k$ è ancora un ideale sinistro proprio di A. Il fatto che I contenga 0, sia chiuso rispetto a inverso additivo e assorba la moltiplicazione sinistra per elementi di A è immediato. I è allora un ideale sinistro se è chiuso rispetto all'operazione di somma. In effetti, se $x, y \in I$, allora esistono $k, l \in K$ tali che $x \in I_k, y \in I_l$; poiché C è una catena rispetto alla relazioni di inclusione, uno tra gli ideali I_k e I_l contiene l'altro e di conseguenza entrambi gli elementi x, y. Ma allora contiene anche la loro somma, che quindi giace in I.

E' importante notare come un ideale sinistro proprio non contenga 1. Poiché $1 \notin I_k$ per ogni $k \in K$, l'unione $I = \bigcup_{k\in K} I_k$ non contiene 1 ed è quindi un ideale proprio. □

Il Lemma di Zorn ci garantisce ora che $\mathcal{F}$ contiene elementi massimali: questi sono esattamente gli ideali sinistri massimali di A.

Osservazione 4.2.
- A lezione abbiamo dimostrato questo fatto per anelli commutativi con unità, ma il caso generale non è più difficile.
- E' importante nella dimostrazione che $1 \neq 0$ in un anello con unità. In effetti, l'anello nullo (con un solo elemento) non ha ideali massimali.
- Si può adattare facilmente l'argomento presentato per dimostrare che ogni ideale proprio di A è contenuto in almeno un ideale massimale.

4.2. **Confronto di cardinalità.** A lezione ho utilizzato il Lemma di Zorn per dimostrare che le cardinalità di insiemi sono sempre confrontabili, e per mostrare che $X \times \{0, 1\}$ ha la stessa cardinalità di X non appena X è infinito. Le riporto qui di seguito, assieme alla dimostrazione che $|X| = |X \times X|$ se X è un insieme infinito (lievemente più complicata).

Proposizione 4.3. *Per ogni scelta di X, Y insiemi, esiste un'applicazione iniettiva di X in Y oppure di Y in X.*

Dimostrazione. E' sufficiente costruire un'applicazione invertibile tra X e un sottoinsieme di Y oppure tra Y e un sottoinsieme di X.

Sia $\mathcal{F} = \{(A, B, f) | A \subset X, B \subset Y, f : A \to B \text{ è invertibile}\}$, e poniamo $(A, B, f) \leq (A', B', f')$ se $A \subset A', B \subset B'$ e $f'|_A = f$. Si vede subito che $\leq$ è una relazione d'ordine: la riflessività e la transitività sono ovvie, e la simmetria non presenta grandi difficoltà. Pertanto, $(\mathcal{F}, \leq)$ è un insieme parzialmente ordinato, ed è non vuoto in quanto $(\emptyset, \emptyset, \mathrm{id})$ appartiene ad $\mathcal{F}$.

Vogliamo mostrare che $(\mathcal{F}, \leq)$ soddisfa le ipotesi del Lemma di Zorn, e cioè che ogni catena in $\mathcal{F}$ ammette un maggiorante. Se $C = \{(A_i, B_i, f_i), i \in I\} \subset \mathcal{F}$ è una catena, poniamo $A = \bigcup_{i\in I} A_i, B = \bigcup_{i\in I} B_i$ e $f(a) = f_i(a)$ se $a \in A_i$: f definisce effettivamente un'applicazione $A \to B$, poiché ogni elemento di A appartiene ad A_i per qualche i. Inoltre, se a appartiene sia ad A_i che ad A_j, allora, a meno di scambiare i con j, deve valere l'inclusione $A_i \subset A_j$ e $f_j|A_i = f_i$. In altre parole, f_i ed f_j coincidono su A_i, e pertanto restituiscono lo stesso risultato se calcolate su $a \in A_i \cap A_j$. L'iniettività e la suriettività di f seguono dall'invertibilità di ciascuna delle f_i. L'elemento (A, B, f) è allora un maggiorante di C.

Poiché $\mathcal{F}$ soddisfa le ipotesi del Lemma di Zorn, deve possedere elementi massimali. Si vede facilmente che se $A \neq X, B \neq Y$, l'elemento $(A, B, f) \in \mathcal{F}$ non può essere massimale. □

Proposizione 4.4. *Se X è infinito, allora esiste un'applicazione invertibile $X \to X \times \{0, 1\}$.*

Dimostrazione. Se X è numerabile, sappiamo già farlo. Se X non è numerabile, considero

$$\mathcal{F} = \{(A, f) | A \subset X \text{ è infinito e } f : A \to A \times \{0, 1\} \text{ è invertibile}\}.$$

Allora $\mathcal{F}$ soddisfa le ipotesi del Lemma di Zorn rispetto all'ordinamento ovvio, e se (A, f) è un suo elemento massimale, allora $X \setminus A$ non può essere infinito: scelto $B \subset X \setminus A$ numberabile, non è complicato mettere anche B in corrispondenza biunivoca con $B \times \{0, 1\}$, ed estendere quindi f ad una bigezione $A \cup Y \to (A \cup Y) \times \{0, 1\}$.

Ma allora X differisce da A per un numero finito di elementi, così come anche $X \times \{0,1\}$ da $A \times \{0,1\}$. Pertanto, $|X| = |A| = |A \times \{0,1\}| = |X \times \{0,1\}|$. □

Corollario 4.5. *Se X è un insieme infinito, l'unione disgiunta di un numero finito di copie di X può essere messa in corrispondenza biunivoca con X.*

Proposizione 4.6. *Se X è infinito, allora esiste un'applicazione invertibile $X \to X \times X$.*

Dimostrazione. Se X è numerabile, sappiamo già farlo. Se X è più che numerabile, considero

$$\mathcal{F} = \{(A, f) | A \subset X \text{ è infinito e } f : A \to A \times A \text{ è invertibile}\}.$$

Innanzitutto, $\mathcal{F}$ è non vuoto, in quanto se $A \subset X$ è numerabile esiste sicuramente $f : A \to A \times A$ invertibile, e quindi $(A, f) \in \mathcal{F}$. Definiamo su $\mathcal{F}$ una relazione data da

$$(A, f) \leq (B, g) \qquad \Longleftrightarrow \qquad A \subset B \ \text{ e } \ g|_A = f.$$

Tale relazione è evidentemente un ordinamento parziale su $\mathcal{F}$. Mostriamo allora che $(\mathcal{F}, \leq)$ soddisfa le ipotesi del Lemma di Zorn. Sia $C = \{(A_i, f_i), i \in I\}$ una catena in $\mathcal{F}$. Se poniamo $A = \cup_{i \in I} A_i$, l'applicazione $f : A \to A \times A$, che soddisfa $f(a) = f_i(a)$ se $a \in A_i$, è ben definita poiché se a appartiene sia ad A_i che ad A_j sicuramente f_i e f_j coincidono sugli elementi comuni.

L'applicazione f è inoltre iniettiva: se $f(a) = f(a')$, dove $a, a' \in A$, allora esiste $i \in I$ tale che A_i contenga sia a che a'. Ma allora f coincide con f_i su tali elementi, e f_i è iniettiva.

La suriettività di f è anche immediata. Comunque scegliamo $a, a' \in A$, esiste $i \in I$ tale che $a, a' \in A_i$. Ma allora (a, a') appartiene all'immagine di f_i. In conclusione $f : A \to A \times A$ è invertibile, e quindi (A, f) appartiene a $\mathcal{F}$: è chiaramente un maggiorante di C.

Poiché le ipotesi del Lemma di Zorn sono soddisfatte, $\mathcal{F}$ ammette un elemento massimale (A, f). Se $|A| < |X|$ allora $|X| = |X \setminus A|$ e quindi $|A| < |X \setminus A|$. In altre parole, esiste un sottoinsieme di $X \setminus A$, chiamiamolo B, che può essere messo in corrispondenza biunivoca con A. Possiamo allora estendere $f : A \to A \times A$ ad un'applicazione invertibile $g : A \cup B \to (A \cup B) \times (A \cup B)$ mettendo in corrispondenza biunivoca B con $(A \times B) \cup (B \times A) \cup (B \times B)$. In effetti, il primo insieme è in corrispondenza biunivoca con A, mentre il secondo è unione disgiunta di tre insiemi che sono in corrispondenza biunivoca con $A \times A$, e può quindi essere messo in corrispondenza biunivoca con $A \times A$. Infine, $|A| = |A \times A|$.

Questo mostra che se $|A| < |X|$, (A, f) non può essere massimale in $\mathcal{F}$. Ma allora $|A| = |X|$ da cui $|A \times A| = |X \times X|$ e poiché $|A| = |A \times A|$ concludiamo che $|X| = |X \times X|$. □

DIPARTIMENTO DI MATEMATICA, UNIVERSITÀ DEGLI STUDI DI ROMA – "LA SAPIENZA"
E-mail address: dandrea@mat.uniroma1.it