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; allo stesso modo permette di esibire comportamenti profondamente antiintuitivi: vi ho parlato, a lezione, della possibilità di costruire — con un utilizzo opportuno dell'assioma della scelta — sottoinsiemi di $\mathbb{R}$ non misurabili secondo Lebesgue; vi ho anche raccontato del bizzarro *paradosso di Banach-Tarski*[1], che garantisce la possibilità di ripartire la palla tridimensionale in un numero finito di pezzi, che possono poi essere ruotati e traslati in modo da ricomporre due palle tridimensionali delle stesse dimensioni di quella iniziale!

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 $(X, \leq)$ un insieme parzialmente ordinato non vuoto nel quale ogni catena ha (almeno) un maggiorante. Allora X possiede elementi massimali.

Se credete che ogni insieme parzialmente ordinato debba contenere elementi massimali, pensate all'insieme X i cui elementi sono i sottoinsiemi finiti di $\mathbb{N}$, ordinato rispetto all'inclusione. Chiaramente nessun elemento di X 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 X 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 X i cui elementi sono tutti i sottoinsiemi della forma $\{0, 1, \dots, n\}$:

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

allora C è chiaramente una catena che non ammette alcun maggiorante in X. 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 X.

Nonostante il nome del Lemma di Zorn, noi lo prenderemo come principio da non dimostrare, cioè come assioma. In effetti 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! Dal momento che la dimostrazione dell'Assioma della scelta

[1] che in realtà paradosso non è, essendo una costruzione totalmente lecita che non fornisce alcuna contraddizione, se non con la nostra logica geometrica intuitiva.

a partire dal Lemma di Zorn è facile, mentre il viceversa è un po' più complicato, noi diamo per buono il Lemma di Zorn, e lo utilizziamo ogni volta che ci serve.

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 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[2], 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 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

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

In altre parole, 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 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 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 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 F, a meno che $J = I$. Ma se $(I, f) \in F$, 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 F. E' qui che entra in gioco il Lemma di Zorn.

Il Lemma di Zorn garantisce l'esistenza di elementi massimali in F non appena siamo in grado di mostrare che ogni catena in F ammette un maggiorante. Sia quindi C una catena in 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 F ammette un maggiorante; grazie al Lemma di Zorn, 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

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

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 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. Un esempio fatto a lezione

L'esempio più eclatante di ragionamento lacunoso esposto a lezione è stato finora (vi ricordo che avevamo già mostrato che ogni insieme infinito possiede un sottoinsieme numerabile) il seguente:

Lemma 3.1. *Ogni insieme infinito X ammette una partizione in sottoinsiemi tutti numerabili.*

Dimostrazione. Da ogni insieme infinito X posso sottrarre un sottoinsieme numerabile $Y \subset X$ in modo da ottenere $X \setminus Y$. Posso quindi ripetere lo stesso procedimento su $X \setminus Y$.

Il procedimento si arresta solamente quando quello che mi rimane non è un insieme infinito, ma finito. A questo punto, posso aggiungere il numero finito di elementi che rimangono ad una delle parti numerabili, in quanto l'unione di un insieme numerabile e di uno finito è ancora numerabile. □

Questa dimostrazione può anche sembrare convincente — ed in effetti può essere formalizzata in modo da risultare corretta — ma così com'è non può funzionare.

Col senno di poi, cioè dopo aver sfruttato questo lemma per ricavarne ogni possibile conseguenza, sappiamo che

- Se X è numerabile, il lemma è ovvio, perché mi basta ripartire X in un solo pezzo
- Se X è più che numerabile, e Y è un suo sottoinsieme numerabile, la cardinalità di $X \setminus Y$ è uguale a quella di X! Quindi dopo ogni passo mi ritrovo nella situazione di partenza!!!

In effetti, eseguire un numero finito di passi non porta a nulla, se non ad accantonare pezzi di X sempre più grandi: è qui che viene in aiuto il Lemma di Zorn! Se riuscissi a tener conto *non solo* della cardinalità di quello che rimane, ma anche del sottoinsieme che ho accantonato, assicurandomi l'esistenza di un accantonamento massimale (rispetto a quale ordinamento?), allora probabilmente riuscirei a vincere! Proviamoci:

Dimostrazione corretta del Lemma 3.1. Sia F l'insieme l'insieme delle possibili partizioni in sottoinsiemi tutti numerabili di sottoinsiemi di X, cioè:

$$F = \{(Y, \{U_i, i \in I\}) \mid \quad Y \subset X, \quad \{U_i, i \in I\} \text{ è una partizione di } Y, \quad U_i \text{ è numerabile per ogni } i \in I\}\,.$$

Possiamo definire su F una relazione d'ordine come segue: $(Y, \{U_i, i \in I\}) \leq (Z, \{V_j, j \in J\})$ se e solo se $Y \subset Z$ e per ogni $i \in I$ esiste $j \in J$ tale che $U_i = V_j$. In altre parole, Z contiene Y, e la partizione di Z estende quella di Y.

Un elemento $(Y, \{U_i, i \in I\}) \in F$ è massimale se e solo se non è possibile trovare in $X \setminus Y$ un sottoinsieme numerabile col quale estendere la partizione $\{U_i, i \in I\}$: cioè se $X \setminus Y$ non è un insieme infinito. Se possiamo mostrare l'esistenza di elementi massimali di F, possiamo quindi concludere come nella dimostrazione precedente.

Per mostrare l'esistenza di elementi massimali, procediamo utilizzando il Lemma di Zorn. Una catena in F è una famiglia di sottoinsiemi ognuno contenuto nell'altro, con partizioni in insiemi numerabili l'una contenuta nell'altra. Allora l'unione di tali sottoinsiemi, con la partizione che le estende tutte, è un maggiorante. Ma allora ogni catena contiene un maggiorante, e quindi F possiede elementi massimali. □

Vi lascio qualche esercizio per mettervi alla prova.

Esercizi

(1) Sia $\mathbb{K}$ un campo, e V uno spazio vettoriale su $\mathbb{K}$. Se X è un sottoinsieme di V, una *combinazione lineare* di elementi di X è un'espressione (finita) del tipo $\alpha_1 v_1 + \cdots + \alpha_n v_n$, dove i coefficienti α_i sono elementi di $\mathbb{K}$ e v_i sono elementi distinti di X. Un sottoinsieme $X \subset V$ si dice *linearmente indipendente*, o *libero*, se le uniche combinazioni lineari di elementi di X uguali a 0 sono quelle con tutti i coefficienti nulli. Un sottoinsieme $X \subset V$ *genera* V se ogni elemento di V è uguale ad una combinazione lineare di elementi di X. Una *base* di V è un insieme $X \subset V$ linearmente indipendente che genera V.

Mostrate, utilizzando il Lemma di Zorn, che ogni spazio vettoriale $V \neq \{0\}$ possiede (almeno) una base. *[Considerate l'insieme F i cui elementi sono i sottoinsiemi linearmente indipendenti di V, con la relazione d'ordine parziale data dall'inclusione. Mostrate che ogni catena in F possiede un maggiorante]*

(2) Siano X e Y insiemi. Mostrate, utilizzando il Lemma di Zorn, che esiste un'applicazione iniettiva da X in Y, oppure un'applicazione iniettiva da Y in X. *[Considerate l'insieme F i cui elementi sono le applicazioni invertibili da sottoinsiemi di X in sottoinsiemi di Y. Quale ordinamento va definito su F?]*

(3) (Difficile) Mostrate, utilizzando il Lemma di Zorn, che ogni insieme X può essere dotato di un buon ordinamento. *[Considerate l'insieme F i cui elementi sono coppie $(Y, \preceq)$ dove Y è un sottoinsieme di X e $\preceq$ è un buon ordinamento su Y, e definite su F una relazione d'ordine data da $(Y, \preceq) \leq (Y', \preceq')$ se e solo se: $Y \subseteq Y'$, la relazione $\preceq'$ coincide con $\preceq$ sugli elementi di Y, e $y \preceq' y'$ per ogni $y \in Y$, $y' \in Y' \setminus Y$.]*

Dipartimento di Matematica, Università degli studi di Roma – "La Sapienza"
E-mail address: dandrea@mat.uniroma1.it