Spunti di matematica pura per l'orientamento nelle scuole superiori

Matteo Misurati

24 Novembre 2020

1 Introduzione

L'obiettivo della presentazione è fornire agli studenti di Matematica tre percorsi che possano essere presi come spunto per presentare il corso di studi a ragazzi delle scuole superiori.

Tali argomenti sono stati scelti all'interno dei corsi di matematica pura della triennale, in modo da dare agli studenti un'idea concreta di cosa si trovi all'interno di essi, di come sia fatta la matematica a livello universitario e di incuriosirli su concetti interessanti non trattati dai programmi scolastici.

L'idea nasce da un'esperienza personale di orientamento presso il liceo Ariosto di Ferrara, durante la quale si era deciso di concentrarsi sulle possibilità applicative della matematica, illustrando argomenti quali la crittografia, la dinamica delle folle, il flusso del sangue, matematica finanziaria. Ritengo che affiancare ad essi argomenti di matematica pura, anche solo accennati, possa essere utile nel mostrare ad un ragazzo le peculiarità di questo corso di studi nel panorama delle discipline scientifiche, attirando inoltre quegli studenti a cui piace la matematica pura, non necessariamente finalizzata ad una applicazione, ma al semplice piacere di aumentare la propria conoscenza.

2 Il concetto di gruppo

La moderna teoria dei gruppi nasce nel XIX secolo grazie al lavoro di Évariste Galois, al fine di dimostrare la non esistenza delle formule risolutive per equazioni di grado superiore al 4.

Il concetto di gruppo è alla base di molte strutture algebriche comunemente usate (come anelli, campi e spazi vettoriali). Per questo compare in numerosi campi di studio, quali, ad esempio, la topologia, l'algebra lineare e la crittografia.

Esporremo prima alcuni esempi di gruppi, per poi darne diverse definizioni equivalenti.

Esempio 2.1 *Si consideri l'insieme $\mathbb{Z}$ dei numeri interi. Su di esso è definita l'applicazione addizione, che a due numeri interi associa la loro somma:*

$$\begin{aligned} + : \mathbb{Z} \times \mathbb{Z} &\longrightarrow \mathbb{Z} \\ (x, y) &\longmapsto x + y \end{aligned}$$

Tale applicazione è associativa, ossia si ha che $x + (y + z) = (x + y) + z$, per ogni $x, y, z \in \mathbb{Z}$. Inoltre, 0 è detto elemento neutro dell'addizione, ossia per ogni $x \in \mathbb{Z}$, $0 + z = 0 = z + 0$, ed esiste un inverso bilatero rispetto $a +$ per ogni elemento di $\mathbb{Z}$, ossia, per ogni $z \in \mathbb{Z}$, esiste $-z \in \mathbb{Z}$ tale che $z + (-z) = 0 = (-z) + z$. Tale elemento si chiama "opposto".

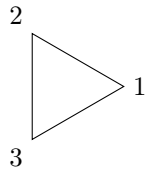
Tali proprietà si ritrovano in altri contesti familiari, come ad esempio nell'insieme dei numeri razionali $\mathbb{Q}$ associato all'applicazione di moltiplicazione:

$$\begin{aligned} \cdot : \mathbb{Q} \times \mathbb{Q} &\longrightarrow \mathbb{Q} \\ (x, y) &\longmapsto x \cdot y \end{aligned}$$

la quale è associativa, ha elemento neutro 1 ed inverso bilatero dato dal reciproco, ossia, per ogni $q = \frac{m}{n} \in \mathbb{Q}$, esiste $q^{-1} = \frac{n}{m}$ tale che $q \cdot q^{-1} = 1$.

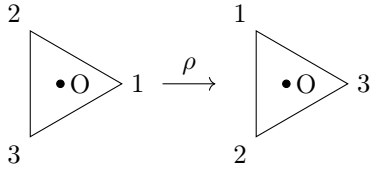
Le proprietà di associatività, esistenza del neutro e dell'inverso, possono però essere trovate anche in situazioni lontane da esempi numerici. Vediamo ora un esempio geometrico.

Le isometrie del piano sono funzioni $f : \mathbb{R}^2 \to \mathbb{R}^2$ che mantengono invariata la distanza tra i punti (ad esempio una rotazione del piano di un certo angolo, attorno ad un certo punto). Consideriamo l'insieme delle isometrie che mandano un poligono in se stesso: sia P_3 il triangolo equilatero in figura e sia D_3 l'insieme delle isometrie $f : \mathbb{R}^2 \to \mathbb{R}^2$ tali che $f(P_3) = P_3$.

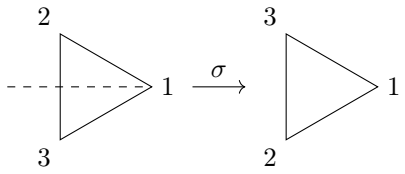


P_3, vertici nominati in senso antiorario.

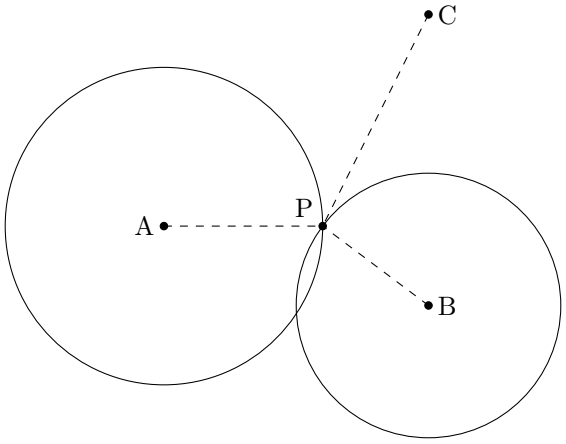
Dimostriamo innanzitutto quali sono gli elementi di D_3. Consideriamone due in particolare, agenti sul triangolo P_3: ρ, corrispondente alla rotazione di angolo $\frac{2\pi}{3}$ del triangolo con centro nell'origine O (situata al centro del triangolo)



e σ, simmetria rispetto alla retta passante per il vertice 1, che è anche asse del lato ad esso opposto.



Mostriamo ora che gli elementi di D_3 sono tutti composizioni di tali isometrie, tenendo presente che un'isometria è unicamente determinata dal valore che assume in tre punti non allineati. Ciò segue dal fatto che, conoscendo la posizione nel piano di tre punti non allineati e la distanza tra ciascuno di questi e un quarto punto P, la posizione di P è univocamente determinata. Infatti, conoscere la distanza di P da due di questi punti significa avere due circonferenze all'interno delle quali è possibile trovare P; esso è quindi contenuto nella loro intersezione: due punti. La distanza dal terzo punto permette di capire quale dei due sia P.



Proposizione 2.2 *Si ha che* $D_3 = \{1, \rho, \rho^2, \sigma, \rho\sigma, \rho^2\sigma\}$*, dove* $\rho^2 = \rho \circ \rho$*,* $\rho^i\sigma := \rho^i \circ \sigma$*, per ogni valore di* i*,* $\circ$ *è il simbolo della composizione di funzioni e* 1 *è l'identità sul piano.*

Dim. Per dimostrare l'uguaglianza tra due insiemi, si mostra come il primo sia contenuto nel secondo, e viceversa.

($\supseteq$) Poiché le isometrie mantengono inalterate le distanze si ha, per ogni $x, y \in \mathbb{R}^2$ e $f, g \in D_3$,

$$|f \circ g(x) - f \circ g(y)| = |f(g(x)) - f(g(y))| = |g(x) - g(y)| = |x - y|$$

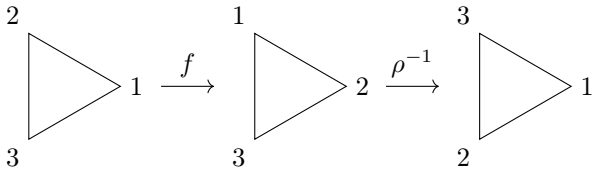
ed inoltre, siccome $f, g \in D_3$ fissano il poligono P_3,

$$f \circ g(P_3) = f(g(P_3)) = f(P_3) = P_3.$$

Ne segue che la composizione di isometrie che fissano P_3 è un'isometria che fissa P_3 e quindi $D_3 \supseteq \{1, \rho, \rho^2, \sigma, \rho\sigma, \rho^2\sigma\}$, essendo ogni elemento di questo insieme o isometria che fissa P_3, o composizioni di isometrie che fissano P_3.

($\subseteq$) Sia ora $f \in D_3$. Essendo un'isometria, che quindi mantiene le distanze fra i punti del piano invariate, l'immagine tramite f di un vertice deve essere ancora un vertice di P_3 (dovendo mantenere una distanza dagli altri vertici pari al lato di P_3) e l'origine viene mandata nell'origine (dovendo mantenere la distanza rispetto ai vertici). Ne segue che f è completamente determinata dal suo comportamento sui vertici 1 e 2, poiché $f((0,0)) = (0,0)$ e un'isometria è univocamente determinata dai valori che assume su tre punti non allineati.

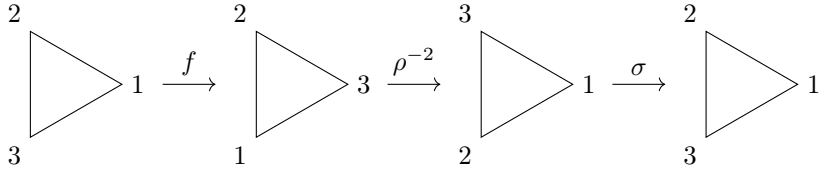
Sia ora $f(1) = i$, con $i \in \{1, 2, 3\}$, e consideriamo l'applicazione $\rho^{-i+1}f$, che consiste nell'applicare f, per poi ruotare il triangolo $(i-1)$-volte di $\frac{\pi}{3}$ in senso orario, ossia di un angolo $-\frac{2\pi(i-1)}{3}$. Si ha $\rho^{-i+1}f(1) = 1$; si veda l'esempio qui sotto per capire meglio come funziona la composizione $\rho^{-i+1}f$ quando $f(1) = i$.



Esempio per $i = 2$.

Essendo vertici adiacenti immagini di vertici adiacenti, ne segue che $\rho^{-i+1}f(2) \in \{2, 3\}$. Se $\rho^{-i+1}f(2) = 2$, si ha che $\rho^{-i+1}f = Id$, essendo ogni elemento di D_3 determinato dal valore che assume in due vertici, quindi $f = \rho^{i-1} \circ \rho^{-i+1}f = \rho^{i-1}$ (ruotiamo il triangolo consecutivamente in direzioni opposte), e $\rho^{-i+1} \in \{1, \rho, \rho^2\} \subseteq \{1, \rho, \rho^2, \sigma, \rho\sigma, \rho^2\sigma\}$.

Se invece $\rho^{-i+1}f(2) = 3$, si ha che $\sigma\rho^{-i+1}f(2) = 2$; ne segue che $\sigma\rho^{-i+1}f = Id$ e quindi $f = \rho^{i-1}\sigma \circ \sigma\rho^{-i+1}f = \rho^{i-1}\sigma$ (applicare consecutivamente la simmetria assiale σ equivale ad applicare l'identità), e $\rho^{-i+1}\sigma \in \{\sigma, \rho\sigma, \rho^2\sigma\} \subseteq \{1, \rho, \rho^2, \sigma, \rho\sigma, \rho^2\sigma\}$.



Esempio per $i = 3$ e $\rho^{-i+1}f(2) = 3$.

Abbiamo quindi mostrato che $D_3 \subseteq \{1, \rho, \rho^2, \sigma, \rho\sigma, \rho^2\sigma\}$, da cui la tesi. □

Si consideri ora la funzione composizione su D_3:

$$\begin{aligned} \circ : D_3 \times D_3 &\longrightarrow D_3 \\ (f, g) &\longmapsto fg := f \circ g \end{aligned}$$

Essendo la composizione di funzioni associativa, tale mappa è associativa. L'identità è elemento neutro rispetto ad essa (chiamata 1, adottando la notazione moltiplicativa). Infine, ogni elemento ha un inverso bilatero: $\rho \circ \rho^2 = \rho^2 \circ \rho = 1$ (ruotare tre volte P_3 equivale a rimanere nella posizione originale), $\sigma \circ \sigma = 1$ e $\sigma\rho^2 \circ \rho\sigma = \rho\sigma \circ \sigma\rho^2 = 1$ (grazie alle due equazioni precedenti).

Più in generale si può dimostrare che, per ogni poligono P_n di n lati, l'insieme D_n delle isometrie che fissano P_n associato alla mappa

$$\begin{aligned} D_n \times D_n &\longrightarrow D_n \\ (f, g) &\longmapsto fg := f \circ g \end{aligned}$$

gode delle stesse proprietà.

Al fine di studiare oggetti apparentemente molto diversi, come quelli visti nei precedenti esempi, dal punto di vista di queste proprietà comuni (associatività, esistenza dell'elemento neutro e dell'inverso bilatero), si definisce il concetto di gruppo. Studiando le proprietà dei gruppi è quindi possibile ottenere informazioni su ciascuno degli oggetti che rispettano la definizione, a prescindere dalle loro caratteristiche peculiari.

Per trattare i gruppi, si usano solitamente due notazioni, dette additiva e moltiplicativa, perché prendono le sembianze degli esempi dati da $(\mathbb{Z}, +)$ e $(\mathbb{Q}, \bullet)$.

Definizione 2.3 *Un gruppo $(G, +)$ è un insieme non vuoto G munito di un'operazione associativa $+ : G \times G \to G$ (in notazione moltiplicativa: $\times : G \times G \to G$) tale che esista un elemento neutro $0 \in G$ (rispettivamente, in notazione moltiplicativa, $1 \in G$) e per ogni elemento x in G esista in G l'inverso bilatero x', tale che $x + x' = x' + x = 0$. Si nota $-x := x'$ (in notazione moltiplicativa $x^{-1} := x'$, se $xx' = x'x = 1$).*

In un gruppo, per ogni elemento, il suo inverso bilatero è unico. Infatti, supponendo che esistano $x', x'' \in G$ inversi bilateri di $x \in G$, si ha che

$$x' = x' + 0 = x' + (x + x'') = (x' + x) + x'' = 0 + x'' = x''.$$

Analogamente, anche l'elemento neutro è unico, poiché, dati $u, u' \in G$ elementi neutri, si ha $u = u + u' = u'$.

Oltre a quella appena data, esistono altre definizioni di gruppo, ad essa equivalenti; ciò significa che essere un gruppo rispetto a una di queste definizioni implica essere un gruppo anche rispetto alle altre. Ne mettiamo in evidenza due attraverso le seguenti proposizioni, usando la notazione additiva.

Proposizione 2.4 *Sia G un insieme non vuoto e sia $+ : G \times G \to G$ un'operazione associativa. Se $u \in G$ è zero destro, ossia $x + u = x$ per ogni $x \in G$, e per ogni $x \in G$ esiste $x' \in G$ inverso destro rispetto a u, ossia $x + x' = u$, allora $(G, +)$ è un gruppo.*

Dim. Dobbiamo dimostrare che u è zero anche a sinistra e che, per ogni $x \in G$, esiste un inverso bilatero $-x$.

Sia x in G, allora esiste il suo inverso destro $x' \in G$ tale che $x + x' = u$. Esiste anche l'inverso destro di x', che chiameremo x'', tale che $x' + x'' = u$. Si ha quindi che

$$\begin{aligned} u + x &= (u + x) + u = u + (x + u) = u + (x + (x' + x'')) \\ &= ((x + x') + x'') = u + (u + x'') = (u + u) + x'' \\ &\overset{\Diamond}{=} u + x'' = (x + x') + x'' = x + (x' + x'') \\ &= x + u = x \end{aligned}$$

ossia u è zero bilatero. Rileggendo ora l'equazione precedente a partire da $\Diamond$, ossia $u + x'' = x$, essendo u uno zero bilatero, si ha $x = x''$, da cui segue

$$x' + x = x' + x'' = 0 = x + x'$$

ossia x' è inverso bilatero di x. □

Considerando zero e inversi sinistri invece che destri, la dimostrazione sarebbe perfettamente analoga. Di conseguenza (G,+), con + associativa, è un gruppo anche se G contiene zero e inversi sinistri.

Proposizione 2.5 *Sia G un insieme non vuoto e sia $+ : G \times G \to G$ un'operazione associativa, tale che, per ogni $a, b \in G$, esistono $x, y \in G$ soluzioni delle equazioni $a + X = b$ e $Y + a = b$. Allora esiste uno zero u in G e $(G, +)$ è un gruppo.*

Dim. G non è vuoto, allora esiste $a_0 \in G$. Consideriamo l'equazione $a_0 + X = a_0$. Sia $u \in G$ la soluzione di tale equazione. Si ha:

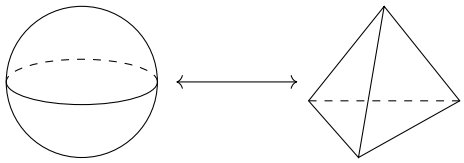
1. Per ogni $g \in G$, abbiamo che $g + u = g$, ossia u è elemento neutro destro di G. Infatti, poiché l'equazione $Y + a_0 = g$ ha una soluzione in G, esiste una $h \in G$ tale che $h + a_0 = g$, da cui otteniamo che $g + u = (h + a_0) + u = h + (a_0 + u) = h + a_0 = g$.

2. Ogni $g \in G$ è invertibile a destra rispetto a u (ossia possiede un inverso destro in G). Infatti, poiché $g + X = u$ ha soluzione, per ogni $g \in G$ esiste $g' \in G$ tale che $g + g' = u$, ossia g è invertibile a destra.

Alla luce di queste proprietà, grazie alla proposizione precedente, abbiamo che (G,+) è un gruppo. □

3 Il concetto di omeomorfismo

La topologia è la parte della matematica che studia quelle proprietà degli oggetti geometrici che rimangono invariate rispetto a trasformazioni continue (informalmente, quelle che permettono di passare da un oggetto ad un altro senza separarne o incollarne parti). Ad esempio, una superficie sferica ed una tetraedrica, ciascuna ottenibile dall'altra con una deformazione continua, pur sembrando molto diverse tra loro, presentano proprietà comuni.

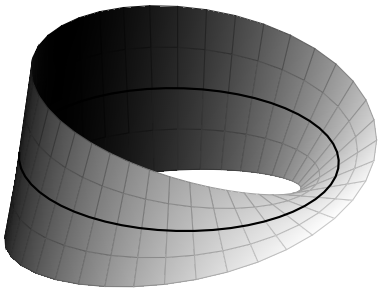


Tra queste:

- entrambe separano lo spazio in due parti, una contenuta dalla superficie, l'altra esterna ad essa;
- entrambe hanno una componente connessa: informalmente, non è possibile scomporre ciascuna di esse in più oggetti senza strappi;
- entrambe sono localmente bidimensionali, ovvero, vicino ad un qualunque punto P sulla superficie, si può fissare un sistema di coordinate rispetto a $\mathbb{R}^2$; ciò significa associare ad ogni punto "vicino" a P una coppia unica di numeri reali che lo identifichi, come viene fatto sul piano attraverso un riferimento cartesiano.

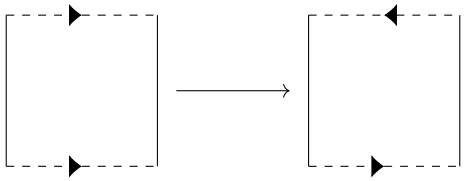
Al fine di risolvere problemi che interessano esclusivamente proprietà come queste, è utile considerare la sfera e il tetraedro come se fossero lo stesso tipo di oggetto. Per far ciò introdurremo il concetto di omeomorfismo; si dirà che la sfera e il tetraedro sono omeomorfi e le proprietà sopracitate verranno dette "invarianti topologiche".

Un esempio classico di oggetti non omeomorfi è dato da una superficie laterale cilindrica e un nastro di Möbius.



Nastro di Möbius, visto in tre dimensioni.

Quest'ultimo, infatti, è ottenibile dalla superficie cilindrica solamente "tagliando" questa lungo l'altezza ed "incollando" le estremità così ottenute nel verso opposto, come mostrato dalla figura seguente. Per questo motivo non possono essere omeomorfi.



A sinistra la superficie laterale cilindrica, a destra il nastro di Möbius, rappresentati nel piano. Le linee tratteggiate sono coincidenti, nel verso indicato dalle frecce.

Inizieremo con l'esporre i concetti di spazio metrico e continuità, che generalizzeremo poi attraverso la definizione di spazio topologico e continuità per funzioni tra spazi topologici. Concluderemo dando una definizione rigorosa di omeomorfismo.

Definizione 3.1 *Uno spazio metrico è una coppia (X, d_X), dove:*

- *X è un insieme.*
- *$d_X : X \times X \to \mathbb{R}^+$ è un'applicazione detta distanza su X, tale che, per ogni $x, y, z \in X$:*
  1. *$d_X(x, y) \geq 0$ e $d_X(x, y) = 0 \Leftrightarrow x = y$*
  2. *$d_X(x, y) = d_X(y, x)$*
  3. *$d_X(x, y) + d_X(y, z) \geq d_X(x, z)$*

Esempio 3.2 *Un semplice esempio di spazio metrico viene offerto da $(\mathbb{R}, d_{\mathbb{R}})$, dove*

$$\begin{aligned} d_{\mathbb{R}} : \mathbb{R} \times \mathbb{R} &\longrightarrow \mathbb{R}^+ \\ (x, y) &\longmapsto |x - y| \end{aligned}$$

è la distanza euclidea sulla retta. $(\mathbb{R}, d_{\mathbb{R}})$, rispettando le proprietà della precedente definizione, è uno spazio metrico. Infatti, per ogni $x, y, z \in \mathbb{R}$,

1. *$|x - y| \geq 0$, ossia $d_{\mathbb{R}}(x, y) \geq 0$. Inoltre $|x - y| = 0$ se e solo se $x - y = 0$, ossia $x = y$.*
2. *$d_{\mathbb{R}}(x, y) = |x - y| = |-(x - y)| = |y - x| = d_{\mathbb{R}}(y, x)$*
3. *Per ogni $a, b \in \mathbb{R}$, si ha $|a + b| \leq |a| + |b|$; infatti, se $a + b$ è positivo, $|a + b| = a + b \leq |a| + |b|$, se è negativo $|a + b| = -(a + b) = -a + (-b) \leq |a| + |b|$. Ne segue quindi che, dati tre punti $x, y, z \in \mathbb{R}$, si ha $d_{\mathbb{R}}(x, z) = |x - z| = |(x - y) + (y - z)| \leq |x - y| + |y - z| = d_{\mathbb{R}}(x, y) + d_{\mathbb{R}}(y, z)$.*

L'esempio precedente mostra come il concetto di spazio metrico, in particolare di distanza, sia utile a formalizzare l'idea di distanza comunemente intesa. Tuttavia, esistono esempi che, pur nella loro semplicità, mostrano come la definizione sia applicabile anche a contesti più alieni.

Uno di questi è il seguente .

Esempio 3.3 *Si consideri la metrica discreta (X, d_X), dove X è un insieme qualunque e d_X è la mappa*

$$d_X : X \times X \longrightarrow \mathbb{R}^+$$
$$(x, y) \longmapsto d_X(x, y) := \begin{cases} 0 & \text{for } x = y \\ 1 & \text{for } x \neq y \end{cases}.$$

Questa mappa associa a due elementi di un insieme X qualsiasi 0 se gli elementi coincidono, 1 altrimenti. Verifichiamo che (X, d_X) rispetti le condizioni per essere uno spazio metrico. Per ogni $a, b, c \in X$:

1. *$d_X(x, y) \geq 0$, essendo o 1 o 0, assumendo valore 0 se e solo se $x = y$ per definizione.*
2. *Se $x = y$, si ha $d_X(x, y) = 0 = d_X(y, x)$. Se $x \neq y$, otteniamo $d_X(x, y) = 1 = d_X(y, x)$.*
3. *Se $x = z$, $d_X(x, z) = 0 \leq d_X(x, y) + d_X(y, z) \in \{0, 2\}$ (infatti o $y = x = z$, o $y \neq x = z$). Se $x \neq z$, $d_X(x, z) = 1 \leq d_X(x, y) + d_X(y, z) \in \{1, 2\}$ (infatti, poiché $x \neq z$, y è o diverso da x e da z, o uguale a x e diverso da z, o viceversa).*

Nonostante, a causa della sua semplicità, non sia particolarmente utile per lo studio dell'insieme X scelto, è interessante come la metrica discreta mostri l'esistenza di almeno una struttura di spazio metrico diversa da quella che intuitivamente si è portati a immaginare.

Procediamo ora con il definire bolle, continuità e spazi topologici.

Definizione 3.4 *Sia (X, d_X) uno spazio metrico, $\epsilon \in \mathbb{R}^+$ e $x_0 \in X$. Allora*

$$B_\epsilon(x_0) := \{x \in X \mid d_X(x, x_0) < \epsilon\},$$

ossia l'insieme degli elementi di X che hanno distanza attraverso d_X minore di ϵ, è detto bolla di raggio ϵ e centro x_0.

Definizione 3.5 *Siano (X, d_X) e (Y, dy) due spazi metrici e sia*

$$f : X \longrightarrow Y$$

un'applicazione. Definiamo la controimmagine di un sottoinsieme $S \subseteq Y$ attraverso f come

$$f^{\leftarrow} = \{x \in X \mid f(x) \in S\}.$$

f è detta continua in x_0 se, per ogni $\epsilon > 0$, esiste $\delta > 0$ tale che

$$f^{\leftarrow}(B_\epsilon(f(x_0))) \supseteq B_\delta(x_0).$$

f è detta continua in X se è continua in x per ogni $x \in X$.

Osservazione 3.6 *Nel caso $(X, d_X) = (Y, d_Y) = (\mathbb{R}, d_{\mathbb{R}})$, la definizione di continuità coincide con quella usata in analisi per le funzioni reali in una variabile reale. Infatti $f^{\leftarrow}(B_\epsilon(f(x_0))) \supseteq B_\delta(x_0)$ se e solo se, per ogni $x \in B_\delta(x_0)$, ossia per ogni $x \in X$ tale che $|x - x_0| < \delta$, si ha $x \in f^{\leftarrow}(B_\epsilon(f(x_0)))$, cioè $f(x) \in (B_\epsilon(f(x_0)))$, a sua volta equivalente a $|f(x) - f(x_0)| < \epsilon$. In sintesi, si ottiene che, se $f : \mathbb{R} \longrightarrow \mathbb{R}$ è continua, per ogni $x \in \mathbb{R}$*

$$\forall \epsilon > 0 \ \exists \delta > 0 \ \textit{t.c.} \ |x - x_0| < \delta \Rightarrow |f(x) - x_0| < \epsilon.$$

Attraverso le bolle appena definite, è possibile generalizzare il concetto di insiemi aperti e chiusi da $\mathbb{R}$ ad uno spazio metrico qualsiasi.

Definizione 3.7 *Sia (X, d_X) uno spazio metrico e sia $U \subseteq X$ un sottoinsieme. U è detto aperto se, per ogni $x \in U$, esiste $\epsilon > 0$ tale che $B_\epsilon(x) \subseteq U$, ossia se ogni punto di U è il centro di una bolla contenuta in esso. Un sottoinsieme $C \subseteq X$ è detto chiuso se $X \setminus C := \{x \in X \mid x \notin C\}$ è aperto.*

Conseguenza della definizione è che gli insiemi vuoto $\varnothing$ e X sono contemporaneamente aperti e chiusi. Infatti, ogni bolla è per definizione un sottoinsieme di X, mentre $\varnothing$ non possiede elementi, quindi verifica la definizione di aperto grazie alle proprietà dell'implicazione logica "$\Rightarrow$" (se l'ipotesi è falsa, nel nostro caso $x \in \varnothing$, l'implicazione è verificata in ogni caso). Mostriamo ora una definizione alternativa di continuità, che ci sarà successivamente utile.

Proposizione 3.8 *Siano (X, d_X) e (Y, d_Y) spazi metrici, sia $f : X \to Y$ un'applicazione. f è continua se e solo se, per ogni $V \subseteq Y$ aperto, $f^{\leftarrow}(V)$ è aperto in X.*

Dim. Sia f continua. Dimostriamo che $f^{\leftarrow}(V)$ è aperto per ogni $V \subseteq Y$ aperto. Se $V \cap Im(f) = \varnothing$, si ottiene $f^{\leftarrow}(V) = \varnothing$, aperto di X. Se invece $V \cap Im(f) \neq \varnothing$, per ogni $x_0 \in f^{\leftarrow}(V)$, si ha $f(x_0) \in V$ per definizione. Essendo V aperto, esiste $\epsilon > 0$ tale che $B_\epsilon(x_0) \subseteq V$. Poiché f è continua, esiste $\delta > 0$ tale che $B_\delta(x_0) \subseteq f^{\leftarrow}(B_\epsilon(f(x_0)))$. Per generalità di x_0 si ha quindi che $f^{\leftarrow}(V)$ è aperto.

Sia ora verificato che, per ogni $V \subseteq Y$ aperto, $f^{\leftarrow}(V)$ è aperto in X. Dobbiamo dimostrare che f è continua. Sia $x_0 \in X$ ed $\epsilon > 0$. Poiché $B_\epsilon(f(x_0))$ è una bolla in Y, è in particolare un aperto, ma allora esiste $\delta > 0$ tale che $B_\delta(x_0) \subseteq f^{\leftarrow}(B_\epsilon(f(x_0)))$, ossia f è continua. □

Consideriamo ora alcune proprietà dell'insieme degli aperti di (X, d_X), che chiameremo $\mathcal{U}_X \subseteq \mathcal{P}(X) := \{S \text{ insieme} \mid S \subseteq X\}$ (ossia $\mathcal{U}_X$ è sottoinsieme dell'insieme delle parti di X).

1. $X, \varnothing \subseteq \mathcal{U}_X$.
2. $A_1, A_2 \in \mathcal{U}_X \Rightarrow A_1 \cap A_2 \in \mathcal{U}_X$.

   Infatti, se $x \in A_1 \cap A_2$, esistono $\epsilon_1, \epsilon_2 > 0$ tali che $B(x)_{\epsilon_1} \subseteq A_1$ e $B(x)_{\epsilon_2} \subseteq A_2$. Quindi $B(x)_{\min\{\epsilon_1, \epsilon_2\}} \subseteq B(x)_{\epsilon_1} \cap B(x)_{\epsilon_2} \subseteq A_1 \cap A_2$.

3. $\{A_j\}_{j\in J} \subseteq \mathcal{U}_X \Rightarrow \bigcup_{j\in J} A_j \in \mathcal{U}_X$.

   Ovvero l'unione di una quantità anche infinita, ma numerabile, di aperti è un aperto. Infatti, ogni elemento dell'unione è contenuto in un aperto della famiglia generatrice $\{A_j\}_{j\in J}$, che quindi contiene a sua volta una bolla, inclusa anche nell'unione.

Le definizioni introdotte fino ad ora sarebbero sufficienti per trattare relazioni tra oggetti geometrici, come quelli descritti negli esempi introduttivi. Tuttavia, come abbiamo visto, per definire la continuità, le bolle, quindi la struttura di spazio metrico, non sono necessarie; è infatti sufficiente definire gli aperti sull'insieme considerato per poter avere un oggetto sul quale abbia senso parlare di continuità.

Generalizzare gli spazi metrici, arrivando a definire gli spazi topologici, è ciò che permette alla topologia di utilizzare i propri strumenti per risolvere anche problemi apparentemente non geometrici.

Definizione 3.9 *Sia X un insieme e $\mathcal{U} \subseteq \mathcal{P}(X)$ un sottoinsieme dell'insieme delle parti di X. $(X,\mathcal{U})$ è detto spazio topologico se:*

1. $X, \varnothing \in \mathcal{U}$,
2. $U_1, U_2 \in \mathcal{U} \Rightarrow U_1 \cap U_2 \in \mathcal{U}$,
3. $\{U_j\}_{\in J} \subseteq \mathcal{U} \Rightarrow \bigcup_{j\in J} U_j \in \mathcal{U}$.

$\mathcal{U}$ è detto topologia su X, mentre $U \subseteq \mathcal{U}$ è detto aperto dello spazio topologico.

Siano ora $(X,\mathcal{U})$ e $(Y,\mathcal{V})$ spazi topologici e $f : X \longrightarrow Y$ una applicazione. f è detta continua se, per ogni $V \in \mathcal{V}$, si ha $f^{\leftarrow}(V) \in \mathcal{U}$.

Considerando quanto visto in precedenza, si ha che ogni spazio metrico è in particolare uno spazio topologico. Inoltre, poiché un sottoinsieme qualsiasi di uno spazio metrico è esso stesso uno spazio metrico (ereditando dall'insieme in cui è contenuto la distanza), esso è anche spazio topologico (ad esempio un cubo in $(\mathbb{R}^3, d_{\mathbb{R}^3})$, o una circonferenza nel piano $(\mathbb{R}^2, d_{\mathbb{R}^2})$).

Possiamo ora definire la relazione di omeomorfismo tra due spazi topologici.

Definizione 3.10 *Siano $(X,\mathcal{U}_X)$ e $(Y,\mathcal{U}_Y)$ due spazi topologici. Si ha $X \approx Y$ (X è omeomorfo a Y) se esistono $f : X \to Y$ e $g : Y \to X$ continue, tali che $f \circ g = Id_Y$ e $g \circ f = Id_X$.*

Equivalentemente, X e Y sono omeomorfi se esiste una applicazione continua biettiva $f : X \to Y$ tale che f^{-1} sia continua. f è detta omeomorfismo.

Informalmente, un omeomorfismo è una trasformazione che modifica un oggetto senza tagliare, incollare o comprimere parti. In molti casi, come quello dell'omeomorfismo tra sfera e tetraedro, risulta complicato definire numericamente f e g. Concludiamo quindi dando un ulteriore esempio, al fine di chiarire il concetto di omeomorfismo da un punto di vista intuitivo.

Esempio 3.11 *Raccogliamo le lettere dell'alfabeto latino in classi di equivalenza rispetto alla relazione di omeomorfismo. Una classe di equivalenza contiene tutti e soli gli elementi di un insieme che sono equivalenti rispetto a determinati tipi di relazioni (dette appunto di equivalenza) delle quali "$\approx$" è parte. Per semplicità non teniamo conto delle grazie.*

La prima che consideriamo è

$$\{C, G, I, J, L, M, N, S, U, V, W, Z\}$$

ossia l'insieme formato dalle lettere composte da un solo tratto. Infatti, ciascuna di esse può essere deformata in un segmento (quindi nella I) senza tagliare, incollare o comprimere nulla.

Abbiamo poi

$$\{D, O\},\ \{P, Q\}\ e\ \{A, R\}$$

ciascun elemento delle quali è omeomorfo ad una circonferenza con, rispettivamente, zero, uno e due segmenti uscenti da essa. Tali classi sono distinte proprio perché non possiamo comprimere questi segmenti all'interno della circonferenza.

Ogni elemento di

$$\{H, K\}$$

è omeomorfo ad un segmento, a ogni estremità del quale sono uniti altri due segmenti (a causa del font scelto, ciò è poco evidente per "K").

Gli elementi di

$$\{E, F, Y, T\}\ e\ \{X\}$$

sono omeomorfi a, rispettivamente, tre e quattro segmenti uniti per una estremità.

Si ha infine

$$\{B\}$$

contenente la sola lettera "B", essendo questa omeomorfa a due circonferenze aventi un punto in comune.

4 Principio di induzione e teorema fondamentale dell'aritmetica

Il teorema fondamentale dell'aritmetica, studiato in origine da Euclide e dimostrato con strumenti moderni da Gauss, afferma che ogni numero naturale maggiore di 1 può essere scritto come prodotto di numeri primi in modo unico.

Esso è uno dei motivi per cui 1 non viene considerato un numero primo. Se così non fosse, si perderebbe l'unicità della fattorizzazione per numeri primi, potendo moltiplicare un numero arbitrario di volte per 1 senza cambiarne il valore del prodotto; si dovrebbe quindi escludere 1 nell'enunciato del teorema.

Al fine di dimostrare il teorema fondamentale dell'aritmetica, introdurremo il principio di induzione ed una proprietà fondamentale dei numeri interi, l'identità di Bezout.

Utilizzando il principio di induzione matematica, per provare che una determinata proprietà vale per tutti gli interi $n \geq n_0 \in \mathbb{Z}$, è sufficiente mostrare che tale proprietà è valida per il valore iniziale n_0 (caso iniziale) e che se è valida per un certo valore $n \in \mathbb{Z}$, è valida anche per $n+1$ (passo induttivo).

Un'immagine utile per visualizzare il principio di induzione è quella data da una fila di tessere del domino. La spinta alla prima tessera corrisponde al valere della proprietà per n_0; che al cadere di una tessera segua la caduta della successiva corrisponde il valere della proprietà per $n+1$ quando questa è verificata per n.

Teorema 4.1 *Sia $P(n)$ una proprietà dipendente da $n \in \mathbb{Z}$ e sia dato $n_0 \in \mathbb{Z}$. Se valgono le seguenti condizioni:*

- *$P(n_0)$ è vera;*
- *per ogni $k \in \mathbb{Z}$ tale che $k \geq n_0$ e $P(k)$ è verificata, si ha che $P(k+1)$ è vera;*

allora $P(n)$ è vera per ogni $n \in \mathbb{Z}$ maggiore di n_0.

Dim. Definiamo gli insiemi $X_{n_0} := \{n \in \mathbb{Z} \mid n \geq n_0\}$ e $S := \{n \in X_{n_0} \mid P(n) \text{ è vera}\}$, ossia l'insieme degli interi $n \geq n_0$ che verificano $P(n)$. Poiché per ipotesi $P(n_0)$ è vera, abbiamo che $n_0 \in S$. Per ogni $k \in S$, è verificata $P(k)$, quindi per ipotesi anche $k+1 \in S$.

Supponiamo per assurdo che $S \neq X_{n_0}$ e sia $S' := X_{n_0} \setminus S$. Sia x il più piccolo elemento di S' (S' è inferiormente limitato). Poiché $n_0 \in S$, abbiamo $n_0 \notin S'$, quindi $x \neq n_0$, ma $x \in X_{n_0}$, allora $x > n_0$. Ne segue $x - 1 \geq n_0$, quindi $x - 1 \in X_{n_0}$. Ma x è elemento minimale di S', allora $x - 1 \in X_{n_0} \setminus S' = X_{n_0} \setminus (X_{n_0} \setminus S) = S$. Ne segue, per quanto visto, $x = (x-1) + 1 \in S$, assurdo. Abbiamo quindi che $S = X_{n_0}$, ossia la proprietà $P(n)$ vale per ogni $n \geq n_0$. □

Quanto appena esposto è anche noto come primo principio di induzione. Il secondo principio di induzione matematica ci permette di dimostrare che una data proprietà $P(n)$ dipendente da $n \in \mathbb{Z}$ è vera per tutti gli $n \geq n_0$, semplicemente mostrando che $P(n_0)$ è valida e che, per ogni $n \geq n_0$, se $P(k)$ è vera per tutte le k tali che $n_0 \leq k \leq n$, allora $P(n+1)$ è verificata.

Teorema 4.2 *Sia $P(n)$ una proprietà dipendente da $n \in \mathbb{Z}$ e sia dato $n_0 \in \mathbb{Z}$. Se valgono le seguenti condizioni:*

- *$P(n_0)$ è vera;*
- *per ogni $n \in \mathbb{Z}$ tale che $n \geq n_0$ e che $P(k)$ sia verificata per ogni $n_0 \leq k \leq n$, si ha che $P(k+1)$ è vera;*

allora $P(n)$ è vera per ogni $n \in \mathbb{Z}$ maggiore di n_0.

Dim. Per ogni $n \geq n_0$, definiamo la proprietà $P'(n)$ come segue: $P'(n)$ è vera se e solo se sono contemporaneamente vere tutte le $P(k)$, con $n_0 \leq k \leq n$.

Poiché $P(n_0)$ è vera per ipotesi, allora lo è anche $P'(n_0)$ (vera se e solo se $P(n_0)$ lo è per definizione). Supponiamo ora che $P'(n)$ sia verificata. Allora sono verificate tutte le $P(k)$, con $n_0 \leq k \leq n$; a cui segue per ipotesi induttiva, essendo in particolare verificata $P(n)$, che $P(n+1)$ è vera. Poiché abbiamo verificato le $P(k)$, con $n_0 \leq k \leq n+1$, abbiamo che $P'(n+1)$ è vera per definizione.

Applicando ora il primo principio di induzione a $P'(n)$, otteniamo che, per ogni $n \geq n_0$, $P'(n)$ è vera. In particolare abbiamo che $P(n)$ è verificata. □

Possiamo ora dimostrare una parte del teorema fondamentale dell'aritmetica.

Proposizione 4.3 *Ogni intero $n > 1$ è rappresentabile come prodotto di numeri primi.*

Dim. Si procede sfruttando il secondo principio di induzione su n. Per $n = 2$, essendo 2 primo, la proprietà è verificata. Supponiamo ora che esista la rappresentazione in numeri primi per ogni $k < n$ e dimostriamo che esiste per n. Se n è primo, abbiamo la fattorizzazione $n = n$. Se n non è primo, abbiamo che è composto, ossia esistono $a, b \in \mathbb{N}$ tali che $n = ab$, con a e b diversi da 1 e da n. Poiché $a, b < n$, abbiamo che esiste per entrambi una fattorizzazione in primi. Otteniamo di conseguenza che anche $n = ab$ è fattorizzabile in primi. □

Al fine di dimostrare l'unicità della fattorizzazione, è necessario avere una definizione alternativa di numeri primi, data dal Lemma di Euclide. Procediamo con i preliminari per la sua dimostrazione.

Dati $a, b \in \mathbb{N}$, indichiamo che a divide b con $a \mid b$.

Definizione 4.4 *Dati due interi $a, b \in \mathbb{N}$, il massimo comune divisore tra a e b è un intero $d > 0$ che divide sia a che b e tale che, se $c \mid a$ e $c \mid b$, allora $c \mid d$. In tal caso si scrive $(a, b) := d$. Se $(a, b) = 1$, a e b sono detti primi fra loro.*

Proviamo ora la validità dell'identità di Bezout, risultato fondamentale per la teoria dei numeri, dal quale discendono importanti teoremi quali il teorema cinese del resto. L'incipit della dimostrazione ricorda quella usata per il primo principio di induzione.

Teorema 4.5 *Siano $a, b \in \mathbb{N}$ due interi positivi non nulli, allora il loro massimo comune divisore esiste. Esistono inoltre $u, v \in \mathbb{Z}$ tali che $(a, b) = au + bv$.*

Dim. Sia $S = \{am + bn \mid m, n \in \mathbb{Z},\ am + bn > 0\}$, ossia l'insieme dei numeri naturali ottenuti come combinazione lineare di a e b (quindi moltiplicandoli rispettivamente per degli interi m e n, per poi sommare i prodotti). $S \subseteq \mathbb{N}$ non è vuoto, infatti $|a| = ea + b \cdot 0 > 0$, dove $e = 1$ se $a \geq 0$ e $e = -1$ se $a < 0$. Sia d il minimo di S, che per definizione dell'insieme ha la forma $d = au + bv$, con $u, v \in \mathbb{Z}$.

Dividendo con il resto a per d (divisione euclidea) otteniamo $a = dq + r$, con $0 \leq r < d$. Se r non fosse nullo, avremmo che $r = a - dq = a - (au + bv)q = a(1 - uq) + b(-vq) \in S$, contro la minimalità di d. Ne segue che $d \mid a$.

Ripetiamo ora gli stessi passaggi su b. Otteniamo prima $b = dq' + r'$, con $0 \leq r' < d$, da cui segue, se $r' \neq 0$, $r' = b - dq' = b - (au + bv)q' = b(1 - vq') + a(-uq') \in S$, contro la minimalità di d. Quindi $d \mid b$.

Se $c \mid a$ e $c \mid b$, abbiamo che $c \mid au + bv = d$. Abbiamo quindi che il massimo comune divisore è $(a, b) = au + bv$. □

Procediamo con il Lemma di Euclide, che fornisce la definizione di numero primo prevalente nella matematica moderna. Essa infatti si presta ad essere generalizzata a contesti diversi dai numeri interi.

Teorema 4.6 (Lemma di Euclide) *p è primo se e solo se, per ogni $a, b \in \mathbb{N}$, $p \mid ab$ implica $p \mid a$ o $p \mid b$.*

Dim. Supponiamo p sia primo e divida ab. Se $p \mid a$, abbiamo finito. Se $p \nmid a$, otteniamo che il massimo divisore comune tra p ed a è $(a, p) = 1 = np + ma$, con $n, m \in \mathbb{Z}$. Allora, posto $ab = k \cdot p$ (k esiste poiché $p \mid ab$), si ha $b = npb + mab = npb + mkp = p \cdot (nb + mk)$, ossia $p \mid b$.

Supponiamo ora che $p \mid ab$ implichi $p \mid a$ o $p \mid b$. Se per assurdo avessimo $p = ab$, con a e b diversi da 1 e p, ossia p non primo, si avrebbe in particolare $p \mid ab$ ($p = ab$ divide se stesso). Quindi $p \mid a$ o $p \mid b$, per ipotesi. Se $p \mid a$, esisterebbe k tale che $a = kp$, allora $p = ab = kpb$ e di conseguenza $1 = kb$. Ma l'unica coppia di interi positivi che moltiplicati tra di loro danno 1 è $(1, 1)$, quindi si avrebbe $b = 1$, che contraddice quanto detto in precedenza per a e b.

Nel caso $p \mid a$, posto $a = k'p$, con identici passaggi si arriva a $1 = ak'$, da cui si ottiene $a = 1$ e la medesima contraddizione. Ne segue che p è primo. □

Segue per induzione dal lemma che, se p è primo e $p \mid a_1 \cdot ... \cdot a_n$, esiste un indice i tra 1 e n tale che $p \mid a_i$.

Osservazione 4.7 *In contesti più generali, la definizione comunemente data di numeri primi in $\mathbb{N}$ e quella contenuta nel Lemma di Euclide non sono necessariamente coincidenti, come ad esempio nel seguente insieme*

$$\mathbb{Z}[\sqrt{-5}] := \{n + m\sqrt{-5} \mid n, m \in \mathbb{Z}\}.$$

Infatti, per $2 \in \mathbb{Z}[\sqrt{-5}]$, si ha che 2 è divisibile solo per 1 e se stesso. Ma, nonostante $2 \mid (1 + \sqrt{-5})(1 - \sqrt{-5}) = 1 - (-5) = 6$, 2 non divide ne $(1 + \sqrt{-5})$ ne $(1 - \sqrt{-5})$.

Siamo ora pronti per la dimostrazione del teorema fondamentale dell'aritmetica.

Teorema 4.8 (Teorema fondamentale dell'aritmetica.) *Per ogni intero positivo $n > 1$, esiste ed è unica la fattorizzazione in numeri primi, a meno dell'ordine dei fattori.*

Dim. Avendo già mostrato nella Proposizione 4.3 l'esistenza della fattorizzazione, ci limitiamo a dimostrarne l'unicità.

Procediamo attraverso il secondo principio di induzione. Per $n = 2$, l'unico numero primo minore uguale a 2 è 2 stesso, quindi $2 = 2$ è l'unica fattorizzazione possibile.

Supponiamo ora la fattorizzazione in primi sia unica per ogni k intero tale che $1 < k < n$. Ne dimostriamo l'unicità per n. Siano $p_1 \cdot ... \cdot p_l = q_1 \cdot ... \cdot q_r$ due fattorizzazioni di n. Poiché $p_1 \mid n = q_1 \cdot ... \cdot q_r$ ed è primo, esiste $\lambda \in \{1, ..., r\}$ tale che $p_1 \mid q_\lambda$; tuttavia, essendo q_λ primo, $p_1 = q_\lambda$. Si ha quindi che l'intero positivo $\frac{n}{p_\lambda} < n$ ammette due fattorizzazioni in primi:

$$p_2 \cdot ... \cdot p_l = q_1 \cdot ... \cdot q_{\lambda-1} q_{\lambda+1} \cdot ... \cdot q_r = \frac{n}{p_\lambda}.$$

Per ipotesi induttiva, la fattorizzazione di tale numero è unica. Ne segue che, per ogni indice $i \in \{2, ..., l\}$, esiste un unico indice $j \in \{1, ..., r\} \setminus \{\lambda\}$ tale che $p_i = q_j$. Conseguenza di ciò è l'uguaglianza fra le fattorizzazioni in primi di n (ovvero, per ogni indice $i \in \{1, 2, ..., l\}$ esiste un unico indice $j \in \{1, ..., r\}$ tale che $p_i = q_j$, dove, in particolare, $p_1 = q_\lambda$).

Abbiamo quindi dimostrato per induzione che la fattorizzazione in primi è unica per ogni intero $n > 0$. □

5 Bibliografia

Nella redazione di queste note sono state consultate le seguenti fonti:

- Richard Courant e Herbert Robbins, "What Is Mathematics?", Oxford University Press (1941)
- Gerardo con Diaz, "Mathematical Induction", Harvard University (2013)
- G.H.Miller, "The evolution of group theory", The Mathematics Teacher, National Council of Teachers of Mathematics (1964)

Dimostrazioni e idee sono inoltre state tratte dai seguenti corsi della laurea triennale in Matematica di Unife:

- Algebra (2017/2018), C.Menini e F.Stumbo
- Analisi matematica I (2017/2018), C.Boiti
- Geometria II (2018/2019), M.Mella
- Number theory (2019/2020), F.A.Ellia
- Teoria dei numeri e fondamenti di crittografia (2018/2019), P.Codecà