Indice

1. Insiemi 2
1.1. Definizione di insieme 2
1.2. Corrispondenze e relazioni 5
1.3. Funzioni 7
1.4. Strutture algebriche 10
2. Numeri interi 12
2.1. Sistemi numerici 12
2.2. Divisione 14
2.3. Basi 16
2.4. Teorema Fondamentale dell'Aritmetica 18
2.5. Equazioni Diofantee 21
2.6. Congruenza Modulo n 23
2.7. Congruenze lineari 27
2.8. Funzione di Eulero 31
2.9. Teorema di Fermat-Eulero 33
2.10. Test di primalitá 35
3. Gruppi 38
3.1. Proprietá dei gruppi 38
3.2. Permutazioni 39
3.3. Polinomi su un campo 41
3.4. Radici di un polinomio 45
4. Crittografia 47
4.1. Introduzione alla crittografia 47
4.2. Algoritmo RSA 49
4.3. Firma digitale tramite RSA 51
5. Teoria dei codici 52
5.1. Introduzione alla teoria dei codici 52
5.2. Codici a blocchi 52
5.3. Codici lineari 55

1. Insiemi

1.1. Definizione di insieme

Prende il nome di **insieme** una qualsiasi collezione di oggetti, detti *elementi* o *membri* dell'insieme. In genere, gli insiemi vengono denotati con le lettere maiuscole dell'alfabeto latino, mentre i loro elementi con le lettere minuscole. Per indicare che l'oggetto a é membro dell'insieme A viene usata la notazione $a \in A$, e si dice che a *appartiene* ad A.
Per rappresentare gli elementi che appartengono ad un insieme é possibile sia in maniera **estensionale**, ovvero semplicemente "elencandoli", oppure in maniera **intensionale**, ovvero specificando una certa proprietá che é posseduta da tutti ed i soli elementi di quell'insieme. Formalmente, viene usata questa notazione:

$$\underbrace{A = \{a_1, a_2, a_3, ...\}}_{\text{forma estensionale}} \quad \underbrace{A = \{a : a \text{ possiede la proprietá caratteristica di } A\}}_{\text{forma intensionale}}$$

Esempio 1.1.1: Sia A l'insieme che contiene i colori che possono comparire in un pixel. A puó venire descritto equivalentemente nei due modi:

$$A = \{\text{rosso}, \text{verde}, \text{blu}\} \quad A = \{a : a \text{ é uno dei colori presenti in un pixel}\}$$

Si noti come un insieme possa essere a sua volta trattato come un oggetto, e quindi essere membro di un'altro insieme. Inoltre, non é ammesso che un insieme contenga piú "copie" dello stesso oggetto. Infine, l'ordine in cui gli elementi di un insieme sono disposti non é rilevante.
Dato un insieme A, il numero di elementi che questo contiene é detto **cardinalitá** e si indica con $|A|$. La cardinalitá di un insieme puó essere sia *finita* che *infinita*, pertanto é ammesso che un insieme possa contenere infiniti elementi.
Siano A e B due insiemi. Si dice che B é un **sottoinsieme** di A se ogni membro di B é anche membro di A, e si indica con $B \subseteq A$. Equivalentemente, si dice che A é un **soprainsieme** di B se ogni membro di B é anche membro di A, e si indica con $A \supseteq B$. Formalmente:

$$B \subseteq A \text{ se e solo se } \forall x \in B, x \in A \quad A \supseteq B \text{ se e solo se } \forall x \in B, x \in A$$

Due insiemi A e B sono **uguali** se contengono gli stessi elementi, ovvero se $A \subseteq B$ e $B \subseteq A$, e si indica con $A = B$. Due insiemi A e B sono diversi se esiste almeno un elemento di A che non é contenuto in B oppure se esiste almeno un elemento di B non contenuto in A, e si indica con $A \neq B$. Si noti come non sia ammesso che due insiemi siano uguali e distinti. Ovvero, se per due insiemi A e B vale $A = B$, allora A e B sono lo stesso insieme.
Siano A e B due insiemi. Se B é un sottoinsieme di A ed al contempo non é uguale ad A si dice che B é un **sottoinsieme proprio** di A, e si indica con $B \subset A$. Equivalentemente, se A é un soprainsieme di B ed al contempo non é uguale a B, si dice che A é un **soprainsieme proprio** di B, e si indica con $A \supset B$. Formalmente:

$$B \subset A \text{ se e solo se } \forall x \in B, x \in A \text{ e } B \neq A \quad A \supset B \text{ se e solo se } \forall x \in B, x \in A \text{ e } B \neq A$$

Per indicare che l'insieme B *non* é un sottoinsieme di A viene usata la notazione $B \nsubseteq A$, mentre per indicare che B *non* é un sottoinsieme proprio di A viene usata la notazione $B \not\subset A$. Similmente, per indicare che l'insieme A *non* é un soprainsieme di B viene usata la notazione $A \not\supseteq B$, mentre per indicare che A *non* é un soprainsieme proprio di B viene usata la notazione $A \not\supset B$.

Lemma 1.1.1: Per qualsiasi insieme A valgono: $A \subseteq A$, $A \supseteq A$, $A = A$, $A \not\subset A$, $A \not\supset A$.

Dimostrazione:
1. Per definizione, $A \subseteq A$ se e solo se $\forall x \in A, x \in A$. Essendo $\forall x \in A, x \in A$ una tautologia, si ha $A \subseteq A$;
2. Analoga alla precedente;
3. Dato che $A \subseteq A$ e $A \supseteq A$, si ha $A = A$;

4. Dato che $A \subseteq A$ e $A = A$, si ha $A \not\subset A$;
5. Analoga alla precedente.

□

L'insieme che non contiene alcun elemento viene detto **insieme vuoto**, e si indica con $\emptyset$ oppure con $\{\}$.

Lemma 1.1.2: L'insieme vuoto é sottoinsieme di ogni insieme (compreso di sé stesso).

Dimostrazione: Dato un qualsiasi insieme A, $\emptyset$ é un sottoinsieme di A se ogni membro di $\emptyset$ é anche membro di A. Dato che $\emptyset$ é l'insieme che non ha alcun membro, di fatto rispetta sempre questa definizione, anche nel caso in cui $A = \emptyset$. □

A partire da un insieme A é possibile costruire l'**insieme potenza** di A, o **insieme delle parti** di A, come l'insieme che contiene tutti i sottoinsiemi di A. L'insieme potenza di A viene indicato con $\mathcal{P}(A)$.

Lemma 1.1.3: Per qualsiasi insieme A (compreso $\emptyset$), valgono $\emptyset \in \mathcal{P}(A)$ e $A \in \mathcal{P}(A)$.

Dimostrazione: Dal Lemma 1.1.1 si ha $\emptyset \subseteq A$, mentre dal Lemma 1.1.2 si ha $A \subseteq A$. Avendo definito $\mathcal{P}(A)$ come l'insieme che contiene tutti i sottoinsiemi di A, $\mathcal{P}(A)$ conterrá certamente (almeno) questi due. □

Esempio 1.1.2: Sia $A = \{\text{rosso}, \text{verde}, \text{blu}\}$. Si ha:

$$\mathcal{P}(A) = \{\emptyset, \{\text{rosso}\}, \{\text{verde}\}, \{\text{blu}\}, \{\text{rosso}, \text{verde}\}, \{\text{rosso}, \text{blu}\}, \{\text{verde}, \text{blu}\}, \{\text{rosso}, \text{verde}, \text{blu}\}\}$$

Dati due insiemi A e B, viene detto **unione** di A e di B l'insieme che contiene tutti gli elementi o di A o di B, e si indica con $A \cup B$:

$$A \cup B = \{x : x \in A \vee x \in B\}$$

Si noti come "$\vee$" non vada inteso in senso disgiuntivo. Ovvero, un certo elemento x appartiene ad $A \cup B$ se appartiene ad A, se appartiene a B oppure se appartiene ad entrambi.

Esempio 1.1.3: Siano $A = \{\text{rosso}, \text{verde}, \text{blu}\}$ e $B = \{\text{verde}, \text{giallo}, \text{rosa}, \text{nero}\}$. Si ha:

$$A \cup B = \{\text{rosso}, \text{verde}, \text{blu}, \text{giallo}, \text{rosa}, \text{nero}\}$$

Dati due insiemi A e B, viene detto **intersezione** di A e di B l'insieme che contiene tutti gli elementi di A e di B, e si indica con $A \cap B$:

$$A \cap B = \{x : x \in A \wedge x \in B\}$$

Si noti come "$\wedge$" vada inteso in senso disgiuntivo. Ovvero, un certo elemento x appartiene ad $A \cap B$ se e soltanto se appartiene contemporaneamente sia ad A che a B.
Se l'intersezione di due insiemi é l'insieme vuoto, ovvero se non esiste alcun elemento che sia presente contemporaneamente in entrambi gli insiemi, si dice che tali insiemi sono **disgiunti**.

Esempio 1.1.4: Siano $A = \{\text{rosso}, \text{verde}, \text{blu}\}$ e $B = \{\text{verde}, \text{giallo}, \text{rosa}, \text{nero}\}$. Si ha:

$$A \cap B = \{\text{verde}\}$$

É possibile generalizzare l'unione di k insiemi $A_1, A_2, A_3, ..., A_k$ come l'insieme che contiene tutti gli x che compaiono in almeno uno dei k insiemi:

$$\bigcup_{i=1}^{k} A_i = (...(A_1 \cup (A_2 \cup (A_3 \cup ...)))) \cup A_k = \{x : \exists i \in \{1, 2, ..., k\} : x \in A_i\}$$

Allo stesso modo, é possibile generalizzare l'intersezione di k insiemi $A_1, A_2, A_3, ..., A_k$ come l'insieme che contiene tutti gli x che compaiono in tutti e k gli insiemi:

$$\bigcap_{i=1}^{k} A_i = (...(A_1 \cap (A_2 \cap (A_3 \cap ...)))) \cap A_k = \{x : x \in A_i \forall i \in \{1, 2, ..., k\}\}$$

Lemma 1.1.4: Siano A, B e C tre insiemi. Per la loro unione e la loro intersezione valgono le proprietá:

Commutativa:
- $A \cap B = B \cap A$;
- $A \cup B = B \cup A$.

Associativa:
- $(A \cap B) \cap C = A \cap (B \cap C)$;
- $(A \cup B) \cup C = A \cup (B \cup C)$.

Distributiva:
- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$;
- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

Il risultato viene generalizzato a k insiemi.

Dati due insiemi A e B, viene detta **differenza** di A e B l'insieme che contiene tutti gli elementi di A che non sono contenuti in B, e si indica con $A \setminus B$:

$$A \setminus B = \{x : x \in A \wedge x \notin B\}$$

Siano A e B due insiemi tali per cui $B \subseteq A$. L'insieme $A - B$ viene detto **complemento** di B rispetto ad A, e si indica con $\overline{B}$. Quando é noto dal contesto rispetto a quale insieme un certo insieme viene complementato, questo viene omesso.

Teorema 1.1.1 (Leggi di De Morgan): Siano A e B due sottoinsiemi di un certo insieme U. Si ha:

$$\overline{A \cap B} = \overline{A} \cup \overline{B} \quad \overline{A \cup B} = \overline{A} \cap \overline{B}$$

Il risultato viene generalizzato a k insiemi.

Siano A e B due insiemi. Viene detto **prodotto cartesiano** di A e di B l'insieme costituito da tutte le possibili coppie ordinate costruite a partire dagli elementi di A e di B, e si indica con $A \times B$.

$$A \times B = \{(a, b) : a \in A \wedge b \in B\}$$

Esempio 1.1.5: Siano $A = \{\text{rosso, verde, blu}\}$ e $B = \{\text{verde, giallo, rosa, nero}\}$. Si ha:

$$\begin{aligned} A \times B = \{&(\text{rosso, verde}), (\text{rosso, giallo}), (\text{rosso, rosa}), (\text{rosso, nero}), \\ &(\text{verde, verde}), (\text{verde, giallo}), (\text{verde, rosa}), (\text{verde, nero}), \\ &(\text{blu, verde}), (\text{blu, giallo}), (\text{blu, rosa}), (\text{blu, nero})\} \end{aligned}$$

Il prodotto cartesiano fra due insiemi puó essere generalizzato a k insiemi $A_1, A_2, ..., A_k$ come all'insieme costruito da tutte le possibili k-uple ordinate costruite a partire dagli elementi di ogni A_i per $i = \{1, ..., k\}$:

$$\prod_{i=1}^{k} A_i = A_1 \times A_2 \times ... \times A_k = \{(a_1, a_2, ..., a_k) : a_1 \in A_1 \wedge a_2 \in A_2 \wedge ... \wedge a_k \in A_k\}$$

Nel caso particolare in cui tutti e k gli insiemi $A_1, A_2, ..., A_k$ siano tutti uguali ad un certo insieme A, per indicare il loro prodotto cartesiano si scrive semplicemente A^k.

1.2. Corrispondenze e relazioni

Dati due insiemi A e B, viene detta **corrispondenza** fra A e B un sottoinsieme $\mathcal{R}$ del loro prodotto cartesiano; nel caso particolare in cui $A = B$, viene detta **relazione** su A.
Dato un insieme A ed una relazione $\mathcal{R}$ su A, per indicare che una coppia $(a, b) \in A \times A$ appartiene a $\mathcal{R}$ si usa dire che a é in *relazione* con b e si usa la dicitura $a\mathcal{R}b$.

Esempio 1.2.1: Sia $A = \{\text{rosso}, \text{verde}, \text{blu}\}$. Si ha:

$$A \times A = \{(\text{rosso}, \text{rosso}), (\text{rosso}, \text{verde}), (\text{rosso}, \text{blu}), (\text{verde}, \text{rosso}), \\ (\text{verde}, \text{verde}), (\text{verde}, \text{blu}), (\text{blu}, \text{rosso}), (\text{blu}, \text{verde}), (\text{blu}, \text{blu})\}$$

Una relazione $\mathcal{R}$ su A potrebbe essere:

$$\mathcal{R} \subseteq A \times A = \{(\text{rosso}, \text{rosso}), (\text{rosso}, \text{verde}), (\text{verde}, \text{verde}), (\text{blu}, \text{verde})\}$$

Dato un insieme A ed una relazione $\mathcal{R}$ su di esso, si dice che $\mathcal{R}$ é una relazione:

- **riflessiva** se $\forall a \in A$ si ha $a\mathcal{R}a$;
- **simmetrica** se $\forall a, b \in A$ $a\mathcal{R}b$ implica $b\mathcal{R}a$;
- **transitiva** se $\forall a, b, c \in A$ $a\mathcal{R}b$ e $b\mathcal{R}c$ implicano $a\mathcal{R}c$;
- **antisimmetrica** se $\forall a, b \in A$ $a\mathcal{R}b$ e $b\mathcal{R}a$ implicano $a = b$.

Una relazione puó rientrare in una, piú di una o anche nessuna di queste categorie.

Esempio 1.2.2: Sia $A = \{\text{rosso}, \text{verde}, \text{blu}\}$. Sia:

$$\mathcal{R}_1 = \{(\text{rosso}, \text{rosso}), (\text{verde}, \text{verde}), (\text{rosso}, \text{verde}), (\text{verde}, \text{rosso}), (\text{verde}, \text{blu})\}$$

- Non é riflessiva, perché $(\text{blu}, \text{blu}) \notin \mathcal{R}_1$;
- Non é simmetrica, perché $(\text{verde}, \text{blu}) \in \mathcal{R}_1$ ma $(\text{blu}, \text{verde}) \notin \mathcal{R}_1$;
- Non é transitiva, perché $(\text{rosso}, \text{verde}) \in \mathcal{R}_1$ e $(\text{verde}, \text{blu}) \in \mathcal{R}_1$ ma $(\text{blu}, \text{rosso}) \notin \mathcal{R}_1$;
- Non é antisimmetrica, perché $(\text{rosso}, \text{verde}) \in \mathcal{R}_1$ e $(\text{verde}, \text{rosso}) \in \mathcal{R}_1$ ma $\text{rosso} \neq \text{verde}$.

Sia invece:

$$\mathcal{R}_2 = \{(\text{rosso}, \text{rosso}), (\text{verde}, \text{verde}), (\text{blu}, \text{blu}), (\text{rosso}, \text{verde}), \\ (\text{verde}, \text{rosso}), (\text{verde}, \text{blu}), (\text{blu}, \text{verde}), (\text{rosso}, \text{blu}), (\text{blu}, \text{rosso})\}$$

Tale relazione é riflessiva, simmetrica e transitiva, ma non é antisimmetrica. Sia infine:

$$\mathcal{R}_3 = \{(\text{rosso}, \text{rosso}), (\text{verde}, \text{verde}), (\text{blu}, \text{blu})\}$$

Tale relazione é riflessiva, simmetrica, transitiva e antisimmetrica.

Dato un insieme A, una relazione $\mathcal{R}$ su A che é (almeno) simmetrica, riflessiva e transitiva viene detta **relazione di equivalenza**. Le relazioni di equivalenza vengono anche spesso indicate con il simbolo $\sim$.
Siano A un insieme e $\sim$ una relazione di equivalenza su A. Preso un qualsiasi elemento $a \in A$, si definisce **classe di equivalenza** di a rispetto ad $\sim$ l'insieme:

$$[a]_\sim = \{b : b \in A \wedge b \sim a\}$$

Ovvero, l'insieme che contiene tutti gli elementi di A che sono in relazione con a. Un qualsiasi elemento di una classe di equivalenza viene detto **rappresentante** di tale classe.

Esempio 1.2.3: Sia $A = \{\text{rosso}, \text{verde}, \text{blu}, \text{giallo}\}$. Si consideri la relazione di equivalenza:

$$\begin{aligned}\sim \ = \{&(\text{rosso}, \text{rosso}), (\text{verde}, \text{verde}), (\text{blu}, \text{blu}), (\text{giallo}, \text{giallo}), (\text{rosso}, \text{giallo}),\\ &(\text{giallo}, \text{blu}), (\text{rosso}, \text{blu}), (\text{blu}, \text{rosso}), (\text{giallo}, \text{rosso}), (\text{blu}, \text{giallo})\}\end{aligned}$$

Si hanno le seguenti quattro classi di equivalenza:

$$[\text{verde}]_\sim = \{\text{verde}\} \qquad [\text{blu}]_\sim = [\text{giallo}]_\sim = [\text{rosso}]_\sim = \{\text{rosso}, \text{giallo}, \text{blu}\}$$

Lemma 1.2.1: Per qualsiasi insieme A, per qualsiasi relazione di equivalenza $\sim$ su A e per qualsiasi $a \in A$, si ha $[a]_\sim \neq \emptyset$.

Dimostrazione: Essendo $\sim$ una relazione di equivalenza, deve essere anche riflessiva, ovvero deve valere $a \sim a$. Pertanto, $[a]_\sim$ deve contenere almeno a, e quindi non é un insieme vuoto. □

Lemma 1.2.2: Siano A un insieme non vuoto e $\sim$ una relazione di equivalenza su A. Per ogni $a, b \in A$, si ha $[a]_\sim = [b]_\sim$ oppure $[a]_\sim \cap [b]_\sim = \emptyset$. Ovvero, o le classi di equivalenza di a e di b sono lo stesso insieme o sono due insiemi disgiunti.

Dimostrazione: Si supponga $[a]_\sim \cap [b]_\sim \neq \emptyset$, e sia $c \in [a]_\sim \cap [b]_\sim$. Per definizione di intersezione, si ha $c \in [a]_\sim$ e $c \in [b]_\sim$, ma questo equivale a dire $c \sim a$ e $c \sim b$. Essendo $\sim$ una relazione di equivalenza, deve essere simmetrica, pertanto valendo $c \sim a$ vale anche $a \sim c$. Per lo stesso motivo, deve essere anche transitiva, pertanto valendo $a \sim c$ e $c \sim b$ allora vale anche $a \sim b$, cioè $a \in [b]_\sim$. Essendo $\sim$ simmetrica, se vale $a \sim b$ allora vale anche $b \sim a$.
Sia $x \in A$ un elemento generico per cui vale $x \in [a]_\sim$, ovvero $x \sim a$. Avendo provato che vale $a \sim b$ ed essendo $\sim$ transitiva, vale anche $x \sim b$, ovvero $x \in [b]_\sim$. Essendo x un elemento generico, significa che questa proprietá vale per qualsiasi elemento di $[a]_\sim$, ovvero che qualsiasi elemento di $[a]_\sim$ é anche elemento di $[b]_\sim$. In altre parole, $[a]_\sim \subseteq [b]_\sim$.
Sia $y \in A$ un elemento generico per cui vale $y \in [b]_\sim$, ovvero $y \sim b$. Avendo provato che vale $b \sim a$ ed essendo $\sim$ transitiva, vale anche $y \sim a$, ovvero $y \in [a]_\sim$. Essendo y un elemento generico, significa che questa proprietá vale per qualsiasi elemento di $[b]_\sim$, ovvero che qualsiasi elemento di $[b]_\sim$ é anche elemento di $[a]_\sim$. In altre parole, $[b]_\sim \subseteq [a]_\sim$.
Avendo provato che vale sia $[a]_\sim \subseteq [b]_\sim$ sia $[b]_\sim \subseteq [a]_\sim$, per definizione di uguaglianza fra insiemi vale $[a]_\sim = [b]_\sim$. É stato allora provato che se $[a]_\sim \cap [b]_\sim \neq \emptyset$, allora $[a]_\sim = [b]_\sim$. Ma questa proposizione equivale ad asserire che vale o $[a]_\sim = [b]_\sim$ o $[a]_\sim \cap [b]_\sim = \emptyset$, e pertanto il lemma é provato. □

Dato un insieme A ed una relazione di equivalenza $\sim$ su A, viene detto **insieme quoziente** l'insieme $A/\sim$ che contiene tutte le classi di equivalenza (distinte) di $\sim$. Ovvero:

$$A/\sim \ = \left\{[a]_\sim, a \in A\right\}$$

Esempio 1.2.4: Nell'Esempio 1.2.3 si ha $A/\sim \ = \left\{[\text{blu}]_\sim, [\text{verde}]_\sim\right\}$.

Sia A un insieme diverso da $\emptyset$, e sia $\mathcal{F} = \{X_1, X_2, ..., X_k\}$ un insieme che contiene k sottoinsiemi di A. $\mathcal{F}$ viene detto **partizione** di A se:

- $\forall i \in \{1, ..., k\}$, si ha $X_i \neq \emptyset$;
- $\forall i, j \in \{1, ..., k\}$ si ha $X_i \cap X_j = \emptyset$. Ovvero, ciascun sottoinsieme é disgiunto da tutti gli altri;

- $\bigcup_{i=1}^{k} X_i = A$. Ovvero, l'unione di tutti i sottoinsiemi restituisce l'insieme di partenza.

Esempio 1.2.5: Sia $A = \{\text{rosso, verde, blu, giallo, rosa, nero, bianco, grigio}\}$. Una possibile partizione di tale insieme é data da:

$$\mathcal{F} = \{X_1, X_2, X_3\} = \{\{\text{rosso, nero, bianco, giallo, grigio}\}, \{\text{verde, blu}\}, \{\text{rosa}\}\}$$

Teorema 1.2.1 (Equivalenza fra insieme quoziente e partizioni): Sia A un insieme e sia $\sim$ una relazione di equivalenza su A. L'insieme quoziente $A/\sim$ determina una partizione su A. Allo stesso modo, sia $\mathcal{F} = \{X_1, X_2, ..., X_k\}$ una partizione di A; la relazione $\mathcal{R}$ definita come $a\mathcal{R}b \iff \{\exists i \in \{1, ..., k\} \text{ t.c. } a, b \in X_i\}$ é una relazione di equivalenza su A.

Dimostrazione: Si osservi come:

- Per il Lemma 1.2.1, ogni classe di equivalenza di un qualsiasi insieme non é l'insieme vuoto;
- Per il Lemma 1.2.2, ogni classe di equivalenza di un qualsiasi insieme é o uguale ad un'altra o disgiunta da questa. Essendo l'insieme quoziente costituito da sole classi di equivalenza distinte, si ha che ciascuna classe che lo compone é distinta da tutte le altre;
- Dato che $[a]_\sim \subseteq A$ per qualsiasi $a \in A$, é evidente come $\bigcup_{a\in A} [a]_\sim \subseteq A$. Inoltre, sempre per il Lemma 1.2.1, ogni $a \in A$ appartiene a $[a]_\sim$, e quindi $A \subseteq \bigcup_{a\in A} [a]_\sim$. Unendo questo risultato al precedente, si ha $A = \bigcup_{a\in A} [a]_\sim$.

Ovvero, $A/\sim$ risponde alla definizione di partizione. D'altra parte, sia $\mathcal{R}$ la relazione definita come $a\mathcal{R}b \iff \{\exists i \in \{1, ..., k\} \text{ t.c. } a, b \in X_i\}$. Tale relazione é:

- Riflessiva, perché per definizione di partizione ogni X_i non é vuoto, pertanto esiste sempre almeno un $a \in A$ che vi appartenga, e quindi $a\mathcal{R}a$ é sempre verificato;
- Simmetrica, perché se $a, b \in X_i$ allora $b, a \in X_i$, dato che gli elementi di un insieme non sono ordinati;
- Transitiva, perché se $a, b \in X_i$ e $b, c \in X_i$, allora $a, c \in X_i$.

Pertanto, é una relazione di equivalenza. □

1.3. Funzioni

Siano A e B due insiemi. Una **funzione** (o **applicazione**) da A a B é una legge f che ad ogni elemento di A associa uno ed un solo elemento di B:

$$f : A \mapsto B, f(a) = b$$

Dove A é detto **dominio** di f e B é detto **codominio** di f.
Di fatto, una funzione f da A a B é un caso particolare di una corrispondenza $\mathcal{R}_f$ da A a B dove il secondo termine di ciascuna coppia ordinata che la compone é sempre univoco:

$$f : A \mapsto B \text{ equivale a } \mathcal{R}_f : \forall a \in A, \exists! b = f(a) \in B : (a, b) \in \mathcal{R}_f$$

Per ogni $a \in A$, il suo "corrispettivo" in B, ovvero $b = f(a)$, si dice **immagine** di A. L'insieme che contiene l'immagine di ciascun elemento dell'insieme A, ovvero $f(A) = \{f(a) : a \in A\}$, viene chiamato **immagine** di f, e viene indicato anche semplicemente con $\mathfrak{I}(f)$.
Per ogni $b \in B$, l'elemento a di A per il quale b ne é il "corrispettivo", ovvero $a : f(a) = b$, viene detto **controimmagine** di b. L'insieme che contiene le controimmagini di ciascun elemento dell'insieme B, ovvero $\{a \in A : f(a) \in B\}$, viene chiamato **controimmagine** di f, e viene indicato anche semplicemente con $\mathfrak{I}^{-1}(f)$.

Esempio 1.3.1:
- La legge che associa a ciascun numero razionale $\frac{a}{b}$ associa un numero intero $a+b$ non é una funzione. Questo perché $\frac{a}{b} = \frac{ha}{hb} \forall h \neq 0$, pertanto ad ogni $\frac{a}{b}$ é associata una moltitudine di valori, non uno soltanto. Ad esempio, alla frazione $\frac{2}{3}$ viene associato sia 5, sia 10;
- La legge $f : \mathbb{Z} \to \mathbb{Z}, f(z) = z^2$, che associa a ciascun numero intero il suo quadrato, é una funzione;
- Il sottoinsieme $\{(z, 7), z \in \mathbb{Z}\} \subseteq \mathbb{Z} \times \mathbb{Z}$, ovvero l'insieme composto da tutte le coppie ordinate del prodotto cartesiano di $\mathbb{Z}$ con sé stesso che hanno 7 come secondo elemento, é una funzione. Tale sottoinsieme puó essere scritto in maniera piú esplicita nella forma di legge come $f : \mathbb{Z} \to \mathbb{Z}, f(z) = 7$.

Siano dati due insiemi A e B ed una funzione $f : A \mapsto B$. Si dice che f é **iniettiva** se ad elementi distinti di A vengono sempre associati elementi distinti di B:

$$a_1, a_2 \in A : a_1 \neq a_2 \implies f(a_1) \neq f(a_2)$$

Si dice che f é **suriettiva** se il codominio B e l'insieme $f(A)$ coincidono, ovvero se ogni elemento di B ha almeno una controimmagine:

$$\forall b \in B, \exists a \in A : f(a) = b$$

Si dice che f é **biettiva**, o **biunivoca**, se é sia iniettiva sia suriettiva. In altre parole, f é biettiva se ad elementi distinti di A vengono associati elementi distinti di B e se ciascun elemento di B ha sempre una controimmagine:

$$\forall b \in B, \exists! a \in A : f(a) = b$$

Esempio 1.3.2:
- La funzione $f : \mathbb{Z} \to \mathbb{Z}, f(z) = 0$ non é iniettiva, perché ogni elemento di $\mathbb{Z}$ viene sempre associato allo stesso elemento di $\mathbb{Z}$ (lo 0, in questo caso). Inoltre, non é suriettiva, perché tutti gli elementi del codominio al di fuori di 0 non hanno una controimmagine;
- La funzione $f : \mathbb{Z} \mapsto \mathbb{Z}, f(z) = z^2$ non é iniettiva, perché se per un certo $a \in \mathbb{Z}$ vale $b = f(a)$, anche per $-a \in \mathbb{Z}$ vale $b = f(-a)$. Ad esempio, $f(4) = f(-4) = 16$). Inoltre, non é suriettiva, perché tutti gli elementi di $\mathbb{Z}$ che non sono quadrati perfetti non hanno una controimmagine. Ad esempio, non esiste un $a \in \mathbb{Z}$ tale per cui $f(a) = 13$. Infatti, sebbene esistano due a tali per cui $f(a) = 13$, ovvero $\pm\sqrt{13}$, questi non sono numeri interi, pertanto non appartengono al dominio;
- La funzione $f : \mathbb{N} \mapsto \mathbb{Z}, f(z) = z^2$ é iniettiva, perché ad ogni elemento di $\mathbb{N}$ viene associato un elemento distinto di $\mathbb{Z}$. Non é peró suriettiva, perché tutti gli elementi di $\mathbb{Z}$ che non sono quadrati perfetti non hanno una controimmagine;
- La funzione $f : \mathbb{Z} \mapsto \mathbb{Z}, f(z) = z + 1$ é iniettiva, perché per ogni numero intero esiste uno ed un solo numero intero ottenuto sommandovi uno. É inoltre anche suriettiva, perché per ogni numero intero é sempre possibile trovare un'altro numero intero ottenuto a partire dal precedente avendovi sommato uno. Pertanto, é una funzione biettiva.

Sia A un insieme non vuoto. La funzione $*$ viene detta **operazione binaria** su A se ha come dominio il prodotto cartesiano di A con sé stesso ed il codominio coincidente con A:

$$* : A \times A \mapsto A$$

Esempio 1.3.3:
- La funzione $f : \mathbb{Z} \times \mathbb{N} \mapsto \mathbb{Z}, f(a, b) = a^b$ non é un'operazione binaria;
- La funzione $f : \mathbb{Z} \times \mathbb{Z} \mapsto \mathbb{Z}, f(a, b) = \sqrt[b]{a}$ é un'operazione binaria.

Sia $*$ una operazione su un insieme a, e siano $a, b, c \in A$ tre suoi elementi. Si dice che $*$ gode della **proprietá associativa** se applicare a c il risultato dell'applicazione di $*$ ad a e a b equivale all'applicare ad a il risultato dell'applicazione di $*$ a b e a c. In altri termini:

$$(a * b) * c = a * (b * c)$$

Dove le parentesi tonde determinano l'*ordine di precedenza* dell'applicazione di $*$.
Sia $*$ una operazione su un insieme a, e siano $a, b \in A$ due suoi elementi. Si dice che $*$ gode della **proprietá commutativa** se applicare a a b equivale ad applicare b ad a. In altri termini:

$$a * b = b * a$$

Esempio 1.3.4:
- L'operazione $f : \mathbb{Z} \times \mathbb{Z} \mapsto \mathbb{Z}, f(a, b) = a + b$ gode sia della proprietá associativa che della proprietá commutativa;
- L'operazione $f : \mathbb{Z} \times \mathbb{Z} \mapsto \mathbb{Z}, f(a, b) = a - b$ non gode né della proprietá associativa né della proprietá commutativa;
- L'operazione $f : \mathbb{Z} \times \mathbb{Z} \mapsto \mathbb{Z}, f(a, b) = 2^{a+b}$ gode della proprietá commutativa, ma non di quella associativa. Infatti, sebbene sia vero che $f(a, b) = f(b, a)$ in quanto $2^{a+b} = 2^{b+a}$, non é vero che $f(a, f(b, c)) = f(f(a, b), c)$, in quanto $2^{a+2b+c} \neq 2^{2a+b+c}$.
- L'operazione $f : \mathbb{Z} \times \mathbb{Z} \mapsto \mathbb{Z}, f(a, b) = b$ gode della proprietá associativa, ma non di quella commutativa. Infatti, sebbene valga $f(a, f(b, c)) = f(f(a, b), c) = c$, si ha $f(a, b) = b$ e $f(b, a) = a$, pertanto $f(a, b) \neq f(b, a)$.

Se $*$ é una operazione sull'insieme A, un elemento $e \in A$ viene detto **elemento neutro** per $*$ se, per qualsiasi $a \in A$, vale $e * a = a * e = a$. Se $*$ é una operazione sull'insieme A che ammette elemento neutro e, per ciascun $a \in A$ esiste un $\tilde{a} \in A$ chiamato **opposto** di A tale per cui $a * \tilde{a} = \tilde{a} * a = e$.

Esempio 1.3.5: L'operazione $f : \mathbb{Z} \times \mathbb{Z} \mapsto \mathbb{Z}, f(a, b) = a + b$ ha come elemento neutro $e = 0$. Infatti, preso un $a \in Z$ qualsiasi, si ha $a + 0 = 0 + a = a$. L'elemento opposto di a rispetto a tale operazione é $-a$, in quanto $a + (-a) = (-a) + a = 0$.

Le operazioni binarie possono venire generalizzate con prodotti cartesiani n-dimensionali. La funzione $*$ viene detta **operazione n-aria** su A se ha come dominio A^n e sé stesso come codominio:

$$* : A^n \mapsto A$$

Per un qualsiasi insieme non vuoto A é possibile costruire la **funzione identitá** i_A come la funzione che ad ogni elemento di A associa sé stesso. Formalmente:

$$i_A : A \mapsto A,\ i_A(a) = a\ \forall a \in A$$

Siano A, B, C e D quattro insiemi. Siano poi $f : A \mapsto B$ e $g : C \mapsto D$ due funzioni, dove $\mathfrak{I}(f) \subseteq C$. Viene detta **funzione composta** di f e di g la funzione che si ottiene applicando la funzione g al risultato della funzione f, ovvero:

$$g \circ f : A \mapsto D,\ g(f(a))\ \forall a \in A$$

Teorema 1.3.1: La composizione di funzioni gode della proprietá associativa. Ovvero, Siano A, B, C, D, E e F sei insiemi. Siano poi $f : A \mapsto B$, $g : C \mapsto D$ e $h : E \mapsto F$ tre funzioni, dove $\mathfrak{I}(f) \subseteq C$ e $\mathfrak{I}(g) \subseteq E$. Allora $h \circ (g \circ f) = (h \circ g) \circ f$.

Lemma 1.3.1: Siano A e B due insiemi, e sia $f : A \mapsto B$ una funzione su questi definita. Allora, per qualsiasi f, $i_B \circ f = f$ e $f \circ i_A = f$.

Corollario 1.3.1: Sia A un insieme e sia $f : A \mapsto A$ una funzione. La composizione di funzioni ha nella funzione identitá l'elemento neutro rispetto all'insieme A^A.

Dimostrazione: Per il Lemma 1.3.1, se $f : A \mapsto B$ é una funzione da un insieme A ad un insieme B, allora $i_B \circ f = f$ e $f \circ i_A = f$. Nel caso particolare in cui $A = B$, si ha $f \circ i_A = i_A \circ f = f$. □

Teorema 1.3.2: Siano $f : A \mapsto B$ e $g : B \mapsto C$ due funzioni, e sia $g \circ f$ la funzione composta di tali funzioni. Si ha allora:

1. Se f e g sono iniettive, allora $g \circ f$ é iniettiva;
2. Se f e g sono suriettive, allora $g \circ f$ é suriettiva;
3. Se f e g sono biettive, allora $g \circ f$ é biettiva;

Dati due insiemi A e B ed una funzione $f : A \mapsto B$, si dice **funzione inversa** di f la funzione f^{-1} tale che, per ogni elemento $b \in B$, $f^{-1}(b)$ é quell'unico $a \in A$ tale per cui $f(a) = b$. Se per una funzione f é possibile costruire la funzione inversa f^{-1}, si dice che f é **invertibile**.

Teorema 1.3.3: Una funzione $f : A \mapsto B$ é invertibile se e solo se é biettiva.

Lemma 1.3.2: Sia $f : A \mapsto B$ una funzione invertibile e sia $g : B \mapsto A$ la sua inversa. Allora $g \circ f = i_A$ e $f \circ g = i_B$. Nel caso particolare in cui $A = B$, si ha $f \circ g = g \circ f = i_A$.

Corollario 1.3.2: Sia $f : A \mapsto A$ una funzione invertibile e sia $g : A \mapsto A$ la sua inversa. La composizione di funzioni ha nella funzione inversa l'inverso rispetto all'insieme A^A.

Dimostrazione: Per il Lemma 1.3.2, se $f : A \mapsto B$ é una funzione invertibile e $g : B \mapsto A$ é la sua inversa, allora $g \circ f = i_A$ e $f \circ g = i_B$. Nel caso particolare in cui $A = B$, si ha $f \circ g = g \circ f = i_A$. □

1.4. Strutture algebriche

Un insieme A su cui sono definite n operazioni $f_1, ..., f_n$ prende il nome di **struttura algebrica** e si indica con $(A, f_1, ..., f_n)$. L'insieme A viene detto **insieme sostegno**, o semplicemente **sostegno**, della struttura algebrica.
Dato che, nella maggior parte dei casi, le operazioni $f_1, ..., f_n$ delle strutture algebriche sono operazioni binarie, se non viene diversamente specificato con "operazione" si intende implicitamente "operazione binaria".
La coppia ordinata $(S, *)$, formata dall'insieme S e da una operazione $*$ definita su S, prende il nome di **semigruppo** se $*$ gode della proprietá associativa.
Un semigruppo $(M, *)$ viene detto **monoide** se l'operazione $*$ definita sull'insieme M ammette elemento neutro.
Un monoide $(G, *)$ viene detto **gruppo** se l'operazione $*$ definita sull'insieme G ammette opposto per ogni elemento di G.

Esempio 1.4.1:
- La coppia $(\mathbb{N}, +)$, dove $+$ indica la somma sui numeri interi comunemente intesa, é un semigruppo, perché $+$ gode della proprietá associativa. É anche un monoide, perché $+$ ammette elemento neutro (il numero 0). Non é peró un gruppo;
- La coppia $(\mathbb{Z}, +)$ é, per gli stessi motivi per cui lo é $(\mathbb{N}, +)$, sia un semigruppo che un monoide. É peró anche un gruppo: per ogni $a \in \mathbb{Z}$ esiste sempre un $-a \in \mathbb{Z}$ tale per cui $a + (-a) = (-a) + a = 0$;
- La coppia $(\mathbb{Q}, \cdot)$, dove $\cdot$ indica il prodotto sui numeri razionali comunemente inteso, é un semigruppo, perché $\cdot$ gode della proprietá associativa. É anche un monoide, perché $\cdot$ ammette elemento neutro (il numero 1). Non é peró un gruppo, perché non esiste l'inverso di 0 rispetto a $\cdot$ (richiederebbe di dividere per 0, che non é possibile);
- La coppia $(\mathbb{Q} - \{0\}, \cdot)$ é, per gli stessi motivi per cui lo é $(\mathbb{Q}, \cdot)$, sia un semigruppo che un monoide. É peró anche un gruppo, perché per ogni $a \in \mathbb{Q}$ esiste sempre un $\frac{1}{a} \in \mathbb{Q}$ tale per cui $a \cdot \frac{1}{a} = \frac{1}{a} \cdot a = 1$.

Un semigruppo, un monoide ed un gruppo si dicono, rispettivamente, **semigruppo abeliano**, **monoide abeliano** e **gruppo abeliano** se l'operazione su questi definita gode della proprietá commutativa.
La struttura algebrica $(A, *, \diamond)$ prende il nome di **anello** se sono rispettate le seguenti proprietá:

- $(A, *)$ é un gruppo abeliano;
- $(A, \diamond)$ é un semigruppo;
- L'operazione $\diamond$ gode della **proprietá distributiva** rispetto a $*$, ovvero:

$$a \diamond (b * c) = (a \diamond b) * (a \diamond c) \quad \text{e} \quad (a * b) \diamond c = (a \diamond c) * (b \diamond c) \quad \forall a, b \in A$$

Se $\diamond$ gode inoltre della proprietá commutativa, ovvero se $(A, \diamond)$ é abeliano, allora si dice che $(A, *, \diamond)$ é un **anello commutativo**.
Se $(A, \diamond)$ é un monoide (oltre che un semigruppo), ovvero se esiste per $\diamond$ un elemento neutro, $(A, *, \diamond)$ é un **anello unitario**. Se non diversamente specificato, nel parlare di "anelli" in generale si stará sottointendendo di stare considerando anelli unitari.

Esempio 1.4.2:
- $(\mathbb{Z}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$ e $(\mathbb{Z}_n, +, \cdot)$ sono anelli commutativi;
- $(\mathrm{Mat}(2 \times 2, \mathbb{Q}), +, \cdot)$ é un anello, ma non é commutativo.

L'anello $(A, *, \diamond)$ prende il nome di **campo** se é commutativo, unitario e se $(A - \{0\}, \diamond)$ é un gruppo.

Esempio 1.4.3:
- $(\mathbb{Q}, +, \cdot)$ é un campo;
- $(\mathbb{Z}_n, +, \cdot)$ é un campo solamente se n é un numero primo;
- $(\mathbb{Z}, +, \cdot)$ non é un campo.

2. Numeri interi

2.1. Sistemi numerici

Sia $\mathbb{N}$ un insieme non vuoto, in cui si fissa un elemento detto *zero*, indicato con 0, ed una funzione $+$ da $\mathbb{N}$ in $\mathbb{N}$. Indicata con a^+ l'immagine di a tramite $+$ al variare di $a \in \mathbb{N}$, si dice che a^+ é *elemento successivo*, o *successore*, di a. Si assuma che per l'insieme $\mathbb{N}$ valgano i seguenti assiomi, detti **Assiomi di Peano**:

1. $0 \neq a^+ \ \forall a \in \mathbb{N}$. Ovvero, non esiste alcun elemento di $\mathbb{N}$ avente 0 come successore;
2. La funzione $+$ é iniettiva. Ovvero, non esistono due $a_1, a_2 \in S$ distinti che abbiano uno stesso a^+ come successore;
3. Se $S \subseteq \mathbb{N}, 0 \in S$ e $s^+ \in S \ \forall s \in S$, allora $S = \mathbb{N}$. Ovvero, se S é un sottoinsieme anche improprio di $\mathbb{N}$ che contiene (almeno) 0 e che, per ciascun elemento di S, ne contiene anche l'immagine tramite $+$, allora S e $\mathbb{N}$ sono lo stesso insieme.

L'insieme $\mathbb{N}$ cosí definito prende il nome di **insieme dei numeri naturali**.

Principio 2.1.1 (Principio del buon ordinamento): Sia S un sottoinsieme non vuoto di $\mathbb{Z}$ limitato inferiormente (esiste un $n_0 \in \mathbb{Z}$ tale che $s \geq n_0, \forall s \in S$). Allora S ha minimo, ovvero esiste un $m \in S$ tale che $s \geq m, \forall s \in S$.

Principio 2.1.2 (Principio di induzione): Dato un numero fissato $n_0 \in \mathbb{Z}$, sia $P(n)$ una proposizione dipendente da $n \in \mathbb{Z}$, con $n \geq n_0$. Si supponga che siano verificate le seguenti ipotesi:

- $P(n_0)$ é vera;
- $\forall n$, supponendo che sia vera $P(n)$ é possibile dimostrare che lo sia anche $P(n+1)$.

Allora $P(n)$ é vera $\forall n \in \mathbb{Z}$

Esempio 2.1.1: Si consideri la seguente proposizione, dipendente da n:

$$\sum_{i=1}^{n}(2i-1) = n^2, \forall n \geq 1$$

É possibile applicarvi il principio di induzione ponendo $n_0 = 1$. Nello specifico:

- $P(1)$ é vera. Infatti, $\sum_{i=1}^{1}(2i-1) = (2 \cdot 1) - 1 = 2 - 1 = 1$ e $1^2 = 1$;
- Supponendo che sia vera $P(n)$, si dimostri che é vera $P(n+1)$, ovvero che sia vera $\sum_{i=1}^{n+1}(2i-1) = (n+1)^2$. Si ha:

  $$\sum_{i=1}^{n+1}(2i-1) = (2(n+1)-1) + \sum_{i=1}^{n}(2i-1) = 2n + 1 + \sum_{i=1}^{n}(2i-1) = 2n + 1 + n^2$$

  Che é peró proprio la formula per il calcolo del quadrato di binomio. Pertanto $n^2 + 1 + 2n = (n+1)^2 = \sum_{i=1}^{n+1}(2i-1)$

Essendo verificate entrambe le ipotesi del principio di induzione, si ha che $P(n)$ é vera $\forall n \geq 1$

Il principio di induzione puó essere riespresso in termini diversi.

Principio 2.1.3 (Principio di induzione forte): Dato un numero fissato $n_0 \in \mathbb{Z}$, sia $P(n)$ una proposizione dipendente da $n \in \mathbb{Z}$, con $n \geq n_0$. Si supponga che siano verificate le seguenti ipotesi:

- $P(n_0)$ é vera;
- $\forall m$ tale che $n_0 \leq m < n$, supponendo che sia vera $P(m)$ é possibile dimostrare che lo sia anche $P(n)$.

Allora $P(n)$ é vera $\forall n \in \mathbb{Z}$

L'aggettivo *forte* non sta ad indicare che il principio di induzione forte abbia un maggior potere espressivo del principio di induzione "standard"; indica semplicemente che si basa su una ipotesi (la seconda) piú forte di quella usata dalla formulazione precedente. Infatti, una dimostrazione compiuta mediante una delle due forme del principio di induzione puó essere convertita in una dimostrazione analoga compiuta nell'altra forma.

Teorema 2.1.1: Il principio di induzione, il principio di induzione forte ed il principio del buon ordinamento sono equivalenti.

Dimostrazione: La dimostrazione si compone di tre parti.

1. Assumendo come vero il principio di induzione, si dimostri la validitá del principio di induzione forte. Sia pertanto $P(n)$ una proposizione dipendente da n e sia $n_0 \in \mathbb{Z}$ un valore fissato. Si supponga che siano verificate le seguenti ipotesi:

   - $P(n_0)$ é vera;
   - $\forall m$ tale che $n_0 \leq m < n$, supponendo che sia vera $P(m)$ é possibile dimostrare che lo sia anche $P(n)$. In particolare, dunque, se $P(n-1)$ é vera allora $P(n)$ é vera. Il principio di induzione implica quindi che $P(n)$ é vera per ogni $n \geq n_0$;

2. Assumendo come vero il principio di induzione forte, si dimostri la validitá del principio del buon ordinamento. Sia pertanto $S \subseteq \mathbb{Z}$ un sottoinsieme non nullo dei numeri interi inferiormente limitato da n_0. Si supponga per assurdo il principio del buon ordinamento non sia valido, ovvero che S non ammetta minimo. Si consideri la proposizione $P(n)$ dipendente da n:

   $$P(n) = \textit{Non esiste alcun numero intero minore o uguale ad } n \textit{ che appartenga ad } S$$

   É possibile applicare a $P(n)$ il principio di induzione forte. La prima ipotesi é verificata, perché se n_0 appartenesse ad S, essendone il limite inferiore, allora ne sarebbe necessariamente anche il minimo. Sia dunque n un intero maggiore di n_0. Si assuma allora che $\forall m$ tale che $n_0 \leq m < n$, supponendo che sia vera $P(m)$ é possibile dimostrare che lo sia anche $P(n)$. Si supponga che $P(n)$ sia falsa: esiste allora qualche $t \leq n, t \in S$. Ma questo non é possibile, perché $\forall t \in \mathbb{Z}, n_0 \leq t \leq n$ si suppone $P(t)$ vera, e quindi $t \notin S$. Occorre allora dedurre che S ammetta minimo, e quindi se si assume come valido il principio di induzione forte allora é valido il principio del buon ordinamento.

3. Assumendo come vero il principio del buon ordinamento, si dimostri la validitá del principio di induzione. Dato un numero fissato $n_0 \in \mathbb{Z}$, sia $P(n)$ una proposizione dipendente da $n \in \mathbb{Z}$, con $n \geq n_0$. Si supponga che siano verificate le seguenti ipotesi:

   - $P(n_0)$ é vera;
   - $\forall n$, supponendo che sia vera $P(n)$ é possibile dimostrare che lo sia anche $P(n+1)$.
     Si consideri l'insieme $S \subseteq \mathbb{Z}$ costituito da tutti gli $n \geq n_0$ per i quali $P(n)$ é falsa. Se il principio di induzione fosse verificato, tale insieme dovrebbe essere l'insieme vuoto. Si assuma per assurdo che tale insieme non sia vuoto: per il principio del buon ordinamento tale insieme deve ammettere un minimo, sia questo m, tale per cui $P(m)$ é falsa.
     Dato che l'insieme contiene solo interi n tali per cui $n \geq n_0$ (ma non tutti), dovrá aversi che $m > n_0$, ovvero che $m - 1 \geq n_0$. Ma allora $P(m-1)$ deve essere vera, perché altrimenti si avrebbe $m - 1 \in S$ ed m non sarebbe il minimo di S. Applicando la seconda ipotesi sopra definita, si ha che

$P(m+1-1) = P(m)$ é vera, ma questo é in contraddizione con quanto evidenziato in precedenza. Occorre allora dedurre che se si assume come valido il principio del buon ordinamento, allora é valido il principio di induzione forte.

□

2.2. Divisione

Dati due numeri interi n e m, con $n > m > 0$, l'operazione di **divisione** permette due interi q e r, chiamati rispettivamente *quoziente* e *resto*, tali che il prodotto fra m e q é il multiplo di m che piú si avvicina ad n per difetto ed il resto $r = n - mq$ misura lo scarto.

Teorema 2.2.1: Siano n e m due numeri interi, con $m \neq 0$. Esiste una ed una sola coppia di interi q ed r tali per cui $n = mq + r$ e $0 \leq r < |m|$

Siano a e b due numeri interi. Se esiste $c \in \mathbb{Z}$ tale che $a = bc$, si dice che b *divide* a, oppure analogamente che a é *divisibile* per b. Per indicare che b divide a viene usata la notazione $b \mid a$; se invece b non divide a, si usa la notazione $b \nmid a$. Se b divide a, si dice anche che b é *multiplo* di a. É immediato verificare che, dato $a \in \mathbb{Z}$, sia ± 1 che $\pm a$ sono certamente divisori di a.
Siano $a, b \in \mathbb{Z}$ non entrambi nulli; si dice che $d \in \mathbb{Z}$ é un **Massimo Comun Divisore** tra a e b se sono verificate entrambe le seguenti due condizioni:

1. $d \mid a$ e $d \mid b$. Ovvero, d é divisore sia di a che di b;
2. Se $c \in \mathbb{Z}$ é tale che $c \mid a$ e $c \mid b$, allora $c \mid d$. Ovvero, tutti i divisori di a che sono anche divisori di b sono anche divisori di d.

Teorema 2.2.2: Dati due numeri $a, b \in \mathbb{Z}$ non entrambi nulli, se d e $\tilde{d}$ sono due Massimi Comun Divisori fra a e b allora devono essere uguali in modulo, ovvero deve aversi $d = \pm\tilde{d}$.

Dimostrazione: Essendo d un Massimo Comun Divisore per a e b, deve valere $d \mid a$ e $d \mid b$. Inoltre, deve valere anche che se $c \in \mathbb{Z}$ é tale che $c \mid a$ e $c \mid b$, allora $c \mid d$.
Essendo peró anche $\tilde{d}$ un Massimo Comun Divisore per a e b, deve valere $\tilde{d} \mid a$ e $\tilde{d} \mid b$. Allora é possibile sostituire c con $\tilde{d}$ nella seconda espressione ed ottenere che $\tilde{d} \mid d$.
É peró possibile operare anche in senso contrario: essendo $\tilde{d}$ un Massimo Comun Divisore per a e b, deve valere anche che se $c \in \mathbb{Z}$ é tale che $c \mid a$ e $c \mid b$, allora $c \mid \tilde{d}$, e valendo $d \mid a$ e $d \mid b$ deve aversi che $d \mid \tilde{d}$.
Esistono allora due numeri $h, k \in \mathbb{Z}$ tali per cui $\tilde{d} = hd$ e $d = \tilde{d}$. Ne segue $\tilde{d} = (hk)\tilde{d}$, e quindi $hk = 1$. Deve allora aversi $h = k = 1$ e quindi $d = \tilde{d}$ oppure $h = k = -1$ e quindi $d = -\tilde{d}$. □

Dal teorema si evince immediatamente che se d é un Massimo Comun Divisore positivo di due numeri interi a e b, allora d é univoco. Tale valore viene indicato con $\mathrm{MCD}(a, b)$.

Teorema 2.2.3 (Esistenza ed unicitá del Massimo Comun Divisore): Per una qualsiasi coppia di numeri interi a e b non entrambi nulli esiste sempre ed é univoco $d = \mathrm{MCD}(a, b)$

Dimostrazione: Innanzitutto, é immediato riconoscere che se $d = \mathrm{MCD}(a, b)$, allora é vero anche $d = \mathrm{MCD}(-a, -b)$. É altrettanto immediato riconoscere che $\mathrm{MCD}(a, b) = \mathrm{MCD}(b, a)$ per qualsiasi a, b. Pertanto, senza perdita di generalitá, é possibile assumere che a e b siano numeri naturali con $a \geq b$.
Se $a = 0$ e $b \neq 0$ si verifica facilmente che $\mathrm{MCD}(a, b) = a$; allo stesso modo, se $b = 0$ e $a \neq 0$ si ha $\mathrm{MCD}(a, b) = b$. Si consideri pertanto il caso piú generale in cui $a \neq 0$ e $b \neq 0$. Devono allora esistere un quoziente q_1 ed un resto r_1 tali per cui é possibile eseguire la divisione:

$$a = bq_1 + r_1, 0 \leq r_1 < b$$

Se $r_1 = 0$, allora $\text{MCD}(a, b) = b$, perché $a = bq_1$ é la definizione stessa di $b \mid a$ e q_1 é arbitrario. Se cosí non é, é possibile ripetere l'operazione e risvolgere i calcoli con un nuovo resto ed un nuovo quoziente. Piú in generale:

$$\begin{array}{lll}
(1) & a = bq_1 + r_1 & r_1 \neq 0 \\
(2) & b = r_1 q_2 + r_2 & r_2 \neq 0 \\
(3) & r_1 = r_2 q_3 + r_3 & r_3 \neq 0 \\
 & \cdots & \\
(k-1) & r_{k-3} = r_{k-2} q_{k-1} + r_{k-1} & r_{k-1} \neq 0 \\
(k) & r_{k-2} = r_{k-1} q_k &
\end{array}$$

Il fatto che prima o poi si giunga ad una k-esima iterazione in cui $r_k = 0$ é garantito dal fatto che tale successione é una successione strettamente crescente di numeri non negativi.
L'ultimo resto non nullo, ovvero r_{k-1}, é precisamente $\text{MCD}(a, b)$. Per verificarlo, é sufficiente osservare come questo possegga entrambe le proprietá enunciate nella definizione di Massimo Comun Divisore:

- Alla riga (k) si ha $r_{k-2} = r_{k-1} q_k$, ovvero $r_{k-1} \mid r_{k-2}$. Sostituendo la riga (k) nella riga $(k-1)$ si ha:

  $$r_{k-3} = r_{k-2} q_{k-1} + r_{k-1} = r_{k-1} q_k q_{k-1} + r_{k-1} = r_{k-1}(q_k q_{k-1} + 1)$$

  Ovvero, $r_{k-1} \mid r_{k-3}$ (Si noti come il raccoglimento é ammesso dato che r_{k-1} é definito come non nullo). Risalendo di riga in riga, é facile convincersi che dalla riga (2) si ottiene $r_{k-1} \mid r_1$ e $r_{k-1} \mid b$. Dalla riga (1) segue $r_{k-1} \mid a$. Avendo dimostrato che $r_{k-1} \mid a$ e $r_{k-1} \mid b$, si ha che r_{k-1} possiede la prima proprietá dell'MCD.
- Sia $c \in \mathbb{Z} - \{0\}$. Siano poi $a = c\overline{a}$ e $b = c\overline{b}$. Sostituendo nella riga (1) si ottiene:

  $$a = bq_1 + r_1 \Rightarrow c\overline{a} = c\overline{b}q_1 + r_1 \Rightarrow r_1 = c\overline{a} - c\overline{b}q_1 \Rightarrow r_1 = c\left(\overline{a} - \overline{b}q_1\right)$$

  Da cui si ha $c \mid r_1$. Ponendo $r_1 = c\overline{r_1}$ e sostituendo nella riga (2), si ha:

  $$b = r_1 q_2 + r_2 \Rightarrow c\overline{b} = c\overline{r_1}q_2 + r_2 \Rightarrow r_2 = c\overline{b} - c\overline{r_1}q_2 \Rightarrow r_2 = c\left(\overline{b} - \overline{r_1}q_2\right)$$

  Da cui si ha $c \mid r_2$. Discendendo di riga in riga ed applicando lo stesso procedimento, si arriva fino a $c \mid r_{k-1}$. Ma questo equivale a dire che, per un c numero intero generico, se $c \mid a$ e $c \mid b$, allora $c \mid r_{k-1}$, e quindi r_{k-1} possiede anche la seconda proprietá dell'MCD.

□

La dimostrazione del Teorema 2.2.3 fornisce implicitamente anche un algoritmo per calcolare, a partire da due numeri interi a e b non entrambi nulli, il loro MCD. Tale algoritmo é strutturato come segue:

1. Si calcola qual'é il piú grande intero q tale per cui é possibile moltiplicarlo per b ottenendo un valore inferiore ad a;
2. Si calcola r come differenza fra qb ed a. Se tale valore é nullo, allora q é MCD per a e b, e l'algoritmo termina;
3. b diventa il nuovo a, mentre r diventa il nuovo b. Dopodiché, si torna al punto 1.

Esempio 2.2.1: L'MCD dei numeri $a = 110143$ e $b = 665$ é 19. Infatti:

$$\begin{aligned}
110143 &= 665 \cdot 165 + 418 \\
665 &= 418 \cdot 1 + 247 \\
418 &= 247 \cdot 1 + 171 \\
247 &= 171 \cdot 1 + 76 \\
171 &= 76 \cdot 2 + 19 \\
76 &= 19 \cdot 4
\end{aligned}$$

Teorema 2.2.4 (Identitá di Bézout): Se a e b sono due numeri interi non entrambi nulli, allora esistono due numeri interi x e y tali per cui vale:

$$ax + by = \text{MCD}(a, b)$$

Dimostrazione: Facendo riferimento al Teorema 2.2.3, si consideri la successione di operazioni. In particolare, la riga (1), ovvero $a = bq_1 + r_1$, puó anche essere riscritta come $r_1 = a(1) + b(-q_1)$. Sostituendo nella riga (2), si ha:

$$b = r_1q_2 + r_2 \Rightarrow b = (a - bq_1)q_2 + r_2 \Rightarrow r_2 = b - aq_2 + bq_1q_2 \Rightarrow r_2 = a(-q_2) + b(q_1q_2 + 1)$$

In questo modo, é possibile ciascun resto come combinazione lineare di a e di b. In particolare per il resto r_{k-1}, che é anche l'MCD di a e di b, esisteranno due valori x e y tali per cui é possibile esprimerlo come combinazione lineare di a e b, e quindi $r_{k-1} = \text{MCD}(a, b) = ax + by$. □

La dimostrazione del Teorema 2.2.4 fornisce implicitamente anche un algoritmo per calcolare, a partire da due numeri interi a e b non entrambi nulli, una possibile coppia x, y di interi tali da soddisfare l'identitá per a e b, fintanto che il loro MCD é noto. Tale algoritmo é strutturato come segue:

1. Si esprime r in funzione di a e di b, spostando ques'ultimo a primo membro ed isolando r a secondo membro;
2. Se r é l'MCD di a e di b, l'algoritmo termina, perché le soluzioni particolari cercate sono i coefficienti di a e di b;
3. Si passa alla riga successiva e si ripete il procedimento, esprimendo i due nuovi a e b in funzione dei precedenti. Si noti come questi, ad ogni iterazione, cambiano di segno.

Esempio 2.2.2: L'MCD dei numeri $a = 110143$ e $b = 665$ é 19. Una soluzione particolare che soddisfa l'identitá di Bézout per questa coppia é ricavata di seguito:

$$\begin{aligned}
110143 = 665 \cdot 165 + 418 &\Rightarrow a &&= 165b + 418 &&\Rightarrow a - 165b &&= 418 \\
665 = 418 \cdot 1 + 247 &\Rightarrow b &&= a - 165b + 247 &&\Rightarrow 166b - a &&= 247 \\
418 = 247 \cdot 1 + 171 &\Rightarrow a - 165b &&= 166b - a + 171 &&\Rightarrow 2a - 331b &&= 171 \\
247 = 171 \cdot 1 + 76 &\Rightarrow 166b - a &&= 2a - 331b + 76 &&\Rightarrow 497b - 3a &&= 76 \\
171 = 76 \cdot 2 + 19 &\Rightarrow 2a - 331b &&= 2(497b - 3a) + 19 &&\Rightarrow 8a - 1325b &&= 19
\end{aligned}$$

Se due numeri interi hanno 1 come Massimo Comun Divisore, allora si dice che tali numeri sono **coprimi** o **primi fra di loro**. Tale definizione puó essere riformulata anche rispetto al Teorema 2.2.4.

Lemma 2.2.1: Due numeri interi a e b sono primi fra di loro se e soltanto se esistono due numeri interi x e y tali per cui vale $ax + by = 1$.

Dimostrazione: Il primo verso dell'implicazione deriva direttamente dalla definizione di numeri coprimi. Infatti, due numeri interi a, e b si dicono coprimi se il loro MCD é 1; sostituendolo nell'identitá di Bézout, si ha precisamente $ax + by = 1$.
Ció che manca da dimostrare é il secondo verso, ovvero che se per due numeri interi a e b esistono due numeri interi x e y tali per cui $ax + by = 1$, allora a e b sono coprimi. Si supponga per assurdo che, se esistono x e y, tali per cui $ax + by = 1$, allora a e b non siano coprimi. Questo significa che il loro MCD non é 1, ovvero che $ax + by \neq 1$, ma questo é in contraddizione con l'ipotesi assunta per assurdo. □

2.3. Basi

Teorema 2.3.1 (Esistenza ed unicitá della rappresentazione dei numeri interi in una certa base): Sia b un intero maggiore o uguale a 2. Ogni numero intero n non negativo può essere scritto in uno ed un solo modo nella forma:

$$n = d_k b^k + d_{k-1} b^{k-1} + ... + d_1 b + d_0 \quad \text{con } 0 \leq d_i < b \ \forall i = 0, ..., k \ d_k \neq 0 \text{ per } k > 0$$

Dimostrazione: La dimostrazione prevede di applicare il principio di induzione forte su n. Per $n = 0$ la proposizione é verificata immediatamente. Si assuma allora che la proposizione sia vera per ogni m con $0 \leq m < n$ e la si dimostri per n.
Innanzitutto, si osservi come sia possibile dividere n per b, ottenendo:

$$n = bq + r \quad \text{con } 0 \leq r < b$$

per un certo q ed un certo r. Per la definizione di divisione, si ha $q < n$. Ma allora q é uno degli m per i quali é valida l'ipotesi assunta, ovvero che esiste uno ed un solo modo per scrivere q nella forma:

$$q = c_{k-1} b^{k-1} + c_{k-2} b^{k-2} + ... + c_1 b + c_0$$

Per certi k valori c_i tali per cui $0 \leq c_i < b$. Sostituendo la seconda espressione nella prima, si ha:

$$n = bq + r = b\big(c_{k-1} b^{k-1} + c_{k-2} b^{k-2} + ... + c_1 b + c_0\big) + r = c_{k-1} b^k + c_{k-2} b^{k-1} + ... + c_1 b^2 + c_0 b + r$$

Ponendo $d_k = c_{k-1}, d_{k-1} = c_{k-2}, ..., d_1 = c_0, d_0 = r$, si ha:

$$n = d_k b^k + d_{k-1} b^{k-1} + ... + d_1 b + d_0 \quad \text{con } 0 \leq d_i < b \ \forall i = 0, ..., k$$

Che é l'ipotesi che si voleva dimostrare.
Per quanto riguarda l'unicità di questa scrittura, questa segue dall'unicità di q e di r. □

Dati $b \in \mathbb{Z}$ con $b \geq 2$ e un numero naturale n tale che:

$$n = d_k b^k + d_{k-1} b^{k-1} + ... + d_1 b + d_0 \quad \text{con } 0 \leq d_i < b \ \forall i = 0, ..., k \ d_k \neq 0 \text{ per } k > 0$$

Gli interi $d_0, d_1, ..., d_k$ si dicono le **cifre** di n in **base** b.
Per indicare in quale base n sta venendo espresso, se ne riportano ordinatamente le cifre aggiungendo la base in pedice alla cifra piú a destra. Nel caso in cui il pedice sia assente, si sta sottointendendo che tale numero sta venendo espresso in base 10.
Una base b fa uso di un numero di cifre pari a $b - 1$, partendo da 0; nel caso in cui la base sia maggiore di 10, si usano dei simboli extra per rappresentare le cifre mancanti.
Se é nota la (unica) rappresentazione di un numero intero non negativo in una certa base b, é sempre possibile ricavarne la rappresentazione in base 10 semplicemente svolgendo l'equazione della definizione. Si noti peró come tale equazione possa anche essere riscritta come:

$$\begin{aligned}
n &= d_k b^k + d_{k-1} b^{k-1} + d_{k-2} b^{k-2} + d_{k-3} b^{k-3} + ... + d_1 b + d_0 \\
&= (d_k b + d_{k-1}) b^{k-1} + d_{k-2} b^{k-2} + d_{k-3} b^{k-3} + ... + d_1 b + d_0 \\
&= ((d_k b + d_{k-1}) b + d_{k-2}) b^{k-2} + d_{k-3} b^{k-3} + ... + d_1 b + d_0 \\
&= (((d_k b + d_{k-1}) b + d_{k-2}) b + d_{k-3}) b^{k-3} + ... + d_1 b + d_0 \\
&= ... \\
&= (...(((d_k b + d_{k-1}) b + d_{k-2}) b + d_{k-3}) b^{k-3} + ... + d_1) b + d_0
\end{aligned}$$

Questa forma é nettamente piú convoluta, ma piú semplice da utilizzare per effettuare la conversione. Infatti, sono necessarie solo k moltiplicazioni per b e k addizioni.

Esempio 2.3.1:

$$61405_7 = (((6 \cdot 7 + 1)7 + 4)7 + 0)7 + 5 = ((42 + 1)7 + 4)49 + 5 = (301 + 4)49 + 5 = 14950$$

Per effettuare la conversione inversa, ovvero ricavare la rappresentazione di un numero n in base b a partire dalla sua rappresentazione in base 10, si osservi come le cifre $d_0, d_1, ..., d_k$ di n non siano altro che i resti delle divisioni:

$$\begin{aligned} n &= bq + d_0 \quad 0 \leq d_0 < b \\ q &= q_1 b + d_1 \quad 0 \leq d_1 < b \\ q_1 &= q_2 b + d_2 \quad 0 \leq d_2 < b \\ &\dots \end{aligned}$$

E così via, finchè non si ottiene quoziente nullo.

Esempio 2.3.2:

$$\begin{aligned} 14950 &= 7 \cdot 2135 + 5 \\ 2135 &= 7 \cdot 305 + 0 \\ 305 &= 7 \cdot 43 + 4 \\ 43 &= 7 \cdot 6 + 1 \\ 6 &= 7 \cdot 0 + 6 \end{aligned}$$

Leggendo dal basso verso l'alto, si ha $14950 = 61405_7$

É facile verificare come maggiore é il numero di cifre che la base in cui un numero é espresso ha a disposizione, minore é il numero di cifre necessarie per rappresentarlo. In particolare, il numero di cifre in base b di un intero non negativo n è dato da:

$$k + 1 = \lfloor \log_b(n) \rfloor + 1 = \left\lfloor \frac{\ln(n)}{\ln(b)} \right\rfloor + 1$$

Perché $b^k \leq n < b^{k+1}$

2.4. Teorema Fondamentale dell'Aritmetica

Sia $p \in \mathbb{Z}$, con $p \geq 2$. Il numero intero p si dice **primo** se, per qualsiasi $a, b \in \mathbb{Z}$, $p \mid ab$ implica $p \mid a$ oppure $p \mid b$. Il numero intero p con $p \geq 2$ viene detto **irriducibile** se i suoi divisori sono solo e soltanto $\pm p$ e ± 1. In altre parole, se vale $a \mid p$ con $a \in \mathbb{Z}$, allora $a = \pm p$ oppure $a = \pm 1$.

Teorema 2.4.1: Il numero $p \in \mathbb{Z}$, con $p \geq 2$ é primo se e solo se é irriducibile (ovvero, le due definizioni sono equivalenti).

Dimostrazione:

- Si supponga che p sia un numero primo. Sia $a \in \mathbb{Z}$ un divisore di p, la cui esistenza é garantita per definizione. Deve allora esistere un certo $b \in \mathbb{Z}$ tale per cui $p = ab$; avendosi $p \mid p$ per qualsiasi numero intero, si ha $p \mid ab$. Essendo p un numero primo, per definizione deve aversi $p \mid a$ oppure $p \mid b$:
  - Se $p \mid a$, allora $p = \pm a$, perché avendo scelto a come divisore di p si ha sia $a \mid p$ che $p \mid a$;
  - Se $p \mid b$, allora deve esistere un certo $c \in \mathbb{Z}$ tale per cui $b = pc$. Ma per ipotesi $p = ab$, pertanto $p = a(pc)$, ovvero $\pm 1 = ac$, da cui si ha $a = \pm 1$.

  In entrambi i casi, p risponde alla definizione di numero irriducibile.
- Si supponga che p sia un numero irriducibile. Siano allora $a, b \in \mathbb{Z}$ tali per cui $p \mid ab$; deve allora esistere un certo $q \in \mathbb{Z}$ tale per cui $ab = pq$. Sia $d = \mathrm{MCD}(a, b)$: per definizione, $d \mid p$. Essendo p un numero irriducibile, deve aversi o $d = p$ oppure $d = 1$:
  - Se $d = p$, allora p é uno dei divisori di a, e quindi $p \mid a$;
  - Se $d = 1$, allora esistono due numeri interi x e y tali per cui é valida l'identitá di Bézout, ovvero $1 = ax + py$. Moltiplicando tale identitá per b, si ha $b = abx + pby$, da cui si deduce $p \mid b$.

In entrambi i casi, p risponde alla definizione di numero primo.

□

Un numero non primo (o, equivalentemente, un numero non irriducibile) viene detto **numero composto**.

Lemma 2.4.1 (Lemma di Euclide): Sia p un numero primo. Se p é il divisore del prodotto di $n \geq 2$ numeri interi, allora p é divisore di almeno uno dei fattori.

Dimostrazione: Si applichi il principio di induzione su n. Se $n = 2$, si ha $p \mid ab$ con $a, b \in \mathbb{Z}$, e per definizione $p \mid a$ oppure $p \mid b$.
Si supponga che la proposizione sia vera per n, ovvero che p sia il divisore di almeno uno dei fattori del prodotto $a_1 \cdot a_2 \cdot ... \cdot a_n$, con $a_1, ..., a_n \in \mathbb{Z}$ sapendo che é divisore del prodotto stesso. Si dimostri pertanto che p sia il divisore di almeno uno dei fattori del prodotto $a_1 \cdot a_2 \cdot ... \cdot a_{n+1}$ sapendo che vale $p \mid (a_1 \cdot ... \cdot a_{n+1})$. Sia $b = a_1 \cdot a_2 \cdot ... \cdot a_n$: é possibile allora scrivere $p \mid b \cdot a_{n+1}$. Si ha quindi $p \mid a_{n+1}$ oppure $p \mid b$: se vale $p \mid a_{n+1}$ il lemma é provato immediatamente, mentre se vale $p \mid b$ allora p divide almeno uno dei fattori di b per l'ipotesi induttiva, ed il lemma é provato comunque. □

Si dice che un numero naturale viene **fattorizzato in numeri primi** quando tale numero viene scritto come prodotto di soli numeri primi (non necessariamente distinti). In genere, una fattorizzazione viene espressa raccogliendo a fattor comune i numeri primi per mettere in evidenza la loro molteplicitá. Naturalmente, la fattorizzazione in numeri primi di un numero primo é sé stesso.

Esempio 2.4.1: Il numero 386672 puó venire riscritto come $11 \cdot 13 \cdot 13 \cdot 13 \cdot 2 \cdot 2 \cdot 2 \cdot 2$. Questa é una fattorizzazione in numeri primi, perché 11, 13 e 2 sono numeri primi. Tale fattorizzazione viene in genere scritta come $11 \cdot 13^3 \cdot 2^4$.

Teorema 2.4.2 (Teorema fondamentale dell'aritmetica): Per ogni numero $n \in \mathbb{N}$ tale che $n \geq 2$ esiste uno ed un solo modo per fattorizzarlo in numeri primi (a meno dell'ordine in cui si dispongono i fattori).

Dimostrazione: Per provare l'esistenza della fattorizzazione in numeri primi di n, si proceda per induzione forte su n. Sia $P(n)$ la proposizione *esiste una fattorizzazione in numeri primi per il numero* n, con $n_0 = 2$.
La proposizione $P(n_0)$ é verificata, perché 2 é un numero primo ed é quindi fattorizzabile in numeri primi. Si consideri pertanto la validitá della proposizione $P(n)$ assumendo che questa sia valida per tutti gli m tali per cui $2 \leq m < n$. Se n é un numero primo, allora $P(n)$ é verificata immediatamente; se invece é un numero composto, allora sará certamente scrivibile come prodotto di due interi, siano questi a e b. Si ha allora $n = ab$, con $2 \leq a$ e $b < n$. Essendo sia a che b minori di n, vale per questi l'ipotesi induttiva, ed esiste quindi una fattorizzazione in numeri primi sia per a che per b, siano queste rispettivamente $a_1 \cdot ... \cdot a_h$ e $b_1 \cdot ... \cdot b_k$. É allora possibile fattorizzare n in numeri primi come $(a_1 \cdot ... \cdot a_h) \cdot (b_1 \cdot ... \cdot b_k)$, pertanto (almeno) una fattorizzazione in numeri primi per n esiste.
Per provare l'unicitá della fattorizzazione in numeri primi di n, si proceda nuovamente per induzione forte su n. Sia $P(n)$ la proposizione *esiste una sola fattorizzazione in numeri primi per il numero* n, con $n_0 = 2$. La proposizione $P(n_0)$ é verificata, perché 2 é un numero primo ed é quindi fattorizzabile in numeri primi in un solo modo (sé stesso). Si dimostri quindi che esista un solo modo per fattorizzare in numeri primi n assumendo che esista un solo modo per fattorizzare tutti gli m con $0 \leq m < n$. Dato che almeno una fattorizzazione in numeri primi per n esiste, si supponga $n = p_1 p_2 ... p_s = q_1 q_2 ... q_t$, dove ciascun p_i con $1 \leq i \leq s$ e ciascun q_j con $1 \leq j \leq t$ é un numero primo (non necessariamente distinto dagli altri). Si vuole dimostrare sia che $s = t$, ovvero che entrambe le fattorizzazioni sono costituite dallo stesso numero di elementi, sia che ogni p_i ha un q_j al quale é equivalente, e che quindi le due fattorizzazioni sono equivalenti membro a membro. Poiché $p_1 \mid p_1, p_2, ..., p_s$ si ha che $p_1 \mid q_1 q_2 ... q_t$, e dunque esiste almeno un j con

$1 \leq j \leq t$ per il quale vale $p_1 \mid q_j$. Senza perdita di generalitá, é possibile assumere che il j in questione sia 1 (eventualmente, é sufficiente riordinare i fattori $q_1, ..., q_t$ per fare in modo che sia cosí), ed é quindi possibile assumere che valga $p_1 \mid q_1$. Essendo peró entrambi numeri primi, se ne deduce che $p_1 = q_1$. Ma allora:

$$n = p_1 p_2 ... p_s = q_1 q_2 ... q_t \Rightarrow \cancel{p_1} p_2 ... p_s = \cancel{p_1} q_2 ... q_t \Rightarrow p_2 ... p_s = q_2 ... q_t$$

Che essendo necessariamente entrambe minori di n, vale per queste l'ipotesi induttiva. □

Per calcolare la (univoca) fattorizzazione di un numero primo occorre trovare un numero primo qualsiasi che ne sia un divisore e ripetere il procedimento sul risultato di tale divisione fintanto che é possibile procedere, ovvero fintanto che tale risultato sia diverso da 1.

Esempio 2.4.2:

$$\begin{aligned} 13796146 \div 13 &= 1061242 \\ 1061242 \div 13 &= 81634 \\ 81634 \div 17 &= 4802 \\ 4802 \div 7 &= 686 \\ 686 \div 7 &= 98 \\ 98 \div 7 &= 14 \\ 14 \div 7 &= 2 \\ 2 \div 2 &= 1 \end{aligned}$$

Teorema 2.4.3 (Teorema di Euclide sui numeri primi): Esistono infiniti numeri primi.

Dimostrazione: Si supponga per assurdo che questo non sia vero, e che i numeri primi siano quindi un insieme finito: sia tale insieme $\{p_1, p_2, ..., p_k\}$. Sia $M = 1 + (p_1 \cdot p_2 \cdot ... \cdot p_k)$: essendo 2 il numero primo piú piccolo, si avrá certamente $M \geq 2$. Essendo poi l'insieme $\mathbb{Z}$ chiuso rispetto al prodotto e alla somma, si ha $M \in \mathbb{Z}$. Sono allora valide le ipotesi del Teorema 2.4.2, ed esiste quindi una ed una sola fattorizzazione in numeri primi per M. Se tale fattorizzazione esiste, allora ciascun elemento p_i di tale fattorizzazione deve esserne anche un divisore. Questo peró non é possibile, perché se si avesse $p_i \mid M$ per un qualsiasi $1 \leq i \leq k$ allora si avrebbe anche $p_i \mid 1 = M - (p_1 \cdot p_2 \cdot ... \cdot p_k)$, e non esiste alcun numero che sia divisore di 1. Occorre pertanto assumere che i numeri primi siano infiniti. □

Siano $a, b \in \mathbb{Z}$ non entrambi nulli; si dice che $m \in \mathbb{Z}$ é un **Minimo Comune Multiplo** tra a e b se sono verificate entrambe le seguenti due condizioni:

1. $a \mid m$ e $b \mid m$. Ovvero, sia a che b sono divisori di m;
2. Se $c \in \mathbb{Z}$ é tale che $a \mid c$ e $b \mid c$, allora $m \mid c$. Ovvero, se sia a che b sono divisori di un generico c, allora anche m é divisore di c.

Teorema 2.4.4: Dati due numeri $a, b \in \mathbb{Z}$ non entrambi nulli, se m e $\widetilde{m}$ sono due Minimi Comuni Multipli fra a e b allora devono essere uguali in modulo, ovvero deve aversi $m = \pm\widetilde{m}$.

Dal teorema si evince immediatamente che se m é un Minimo Comune Multiplo positivo di due numeri interi a e b, allora m é univoco. Tale valore viene indicato con $\mathrm{mcm}(a, b)$.

Teorema 2.4.5 (Esistenza ed unicitá del Minimo Comune Multiplo)**:** Per una qualsiasi coppia di numeri interi a e b non entrambi nulli esiste sempre ed é univoco $m = \text{mcm}(a, b)$. In particolare, $\text{mcm}(a, b) = \frac{ab}{\text{MCD}(a,b)}$.

Dimostrazione: Sia $d = \text{MCD}(a, b)$. Siano poi $a = \tilde{a}d$, $b = \tilde{b}d$ e $m = \frac{ab}{d}$. Sostituendo le espressioni di a e b in m, si ha $m = \frac{\tilde{a}d\tilde{b}\cancel{d}}{\cancel{d}} = \tilde{a}\tilde{b}d = a\tilde{b} = b\tilde{a}$, da cui si evince $a \mid m$ e $b \mid m$, provando il primo requisito della definizione di Minimo Comune Multiplo.
Preso un $c \in \mathbb{Z}$ tale per cui $a \mid c$ e $b \mid c$, ossia tale per cui $c = as = bt$ per certi $s, t \in \mathbb{Z}$, si ha $c = \tilde{a}sd = \tilde{b}td$, ovvero $\tilde{a}s = \tilde{b}t$. Poiché $\text{MCD}\left(\tilde{a}, \tilde{b}\right) = 1$, deve aversi $\tilde{a} \mid t$ e $\tilde{b} \mid s$, ovvero deve valere $t = h\tilde{a}$ e $s = k\tilde{b}$ per certi $h, k \in \mathbb{Z}$. Sostituendo $t = h\tilde{a}$ nell'espressione per c, si ha $c = b\tilde{a}h = mh$, da cui si deduce $m \mid c$, provando il secondo requisito della definizione di Minimo Comune Multiplo. □

2.5. Equazioni Diofantee

Viene detta **equazione diofantea** una equazione nella forma:

$$ax + by = c \quad \text{con } a, b, c, x, y \in \mathbb{Z} \text{ e } a, b, c \neq 0$$

Dove a, b, c sono i *termini noti* e x, y sono le *incognite*.
Essendo x e y interi, le *soluzioni* di tale equazione sono tutte e sole le coppie $(x_0, y_0) \in \mathbb{Z} \times \mathbb{Z}$ tali per cui $ax_0 + by_0 = c$.

Esempio 2.5.1: Si consideri l'equazione diofantea $6x + 5y = 3$. Le coppie $(3, -3)$ e $(8, -9)$ sono sue possibili soluzioni.

Teorema 2.5.1 (Condizione necessaria e sufficiente per la solubilitá delle equazioni diofantee)**:** Si consideri l'equazione diofantea $ax + by = c$, con termini noti non nulli $a, b, c \in \mathbb{Z}$ e incognite $x, y \in \mathbb{Z}$. Tale equazione ammette soluzione se e soltanto se $\text{MCD}(a, b) \mid c$.

Dimostrazione: Si supponga che $ax + by = c$ ammetta una certa soluzione $(x_0, y_0) \in \mathbb{Z} \times \mathbb{Z}$. Deve allora valere $ax_0 + by_0 = c$. Valendo $\text{MCD}(a, b) \mid ax_0 + by_0$ si ha $\text{MCD}(a, b) \mid c$. Pertanto, se una equazione diofantea $ax + by = c$ é risolubile, allora $\text{MCD}(a, b) \mid c$.
Viceversa, si supponga che per l'equazione diofantea $ax + by = c$ valga $\text{MCD}(a, b) \mid c$. Questo equivale a dire che vale $c = \text{MCD}(a, b)\tilde{c}$ per un qualche $\tilde{c} \in \mathbb{Z}$. Per l'identità di Bezout esistono certi $s, t \in \mathbb{Z}$ tali per cui $\text{MCD}(a, b) = as + bt$. Sostituendo nell'equazione precedente, si ha $c = (as + bt)\tilde{c} = as\tilde{c} + bt\tilde{c}$. Ponendo $x_0 = s\tilde{c}$ e $y_0 = t\tilde{c}$, si ha $c = ax_0 + by_0$. Essendo $(x_0, y_0) \in \mathbb{Z} \times \mathbb{Z}$, tale coppia é una possibile soluzione per l'equazione. Pertanto, se per l'equazione diofantea $ax + by = c$ vale $\text{MCD}(a, b) \mid c$, allora tale equazione ha (almeno) una soluzione. □

Esempio 2.5.2: Si consideri l'equazione diofantea $74x + 22y = 10$. Ci si chiede se tale equazione ammetta soluzione. Si calcoli pertanto $\text{MCD}(a, b)$:

$$\begin{aligned} 74 &= 22 \cdot 3 + 8 \\ 22 &= 8 \cdot 2 + 6 \\ 8 &= 6 \cdot 1 + 2 \\ 6 &= 2 \cdot 3 \end{aligned}$$

Da cui si ricava $\text{MCD}(74, 22) = 2$. Essendo $2 \mid 10$, si ha che l'equazione ammette soluzione.

Corollario 2.5.1 (Determinare una soluzione particolare di una equazione diofantea): Si consideri l'equazione diofantea risolubile $ax + by = c$, con termini noti non nulli $a, b, c \in \mathbb{Z}$ e incognite $x, y \in \mathbb{Z}$. Una soluzione particolare $(x_0, y_0) \in \mathbb{Z} \times \mathbb{Z}$ di tale equazione puó essere ottenuta dall'identitá di Bézout che ha a e b per termini noti.

Dimostrazione: Sia $ax + by = \text{MCD}(a, b)$ l'identitá di Bézout per a e b. Moltiplicando ambo i membri per un certo $\tilde{c} \in \mathbb{Z}$, si ha $(ax + by)\tilde{c} = ax\tilde{c} + by\tilde{c} = \text{MCD}(a, b)\tilde{c}$. Sostituendo $x\tilde{c} = x_0$, $y\tilde{c} = y_0$ e $\text{MCD}(a, b)\tilde{c} = c$, si ha $ax_0 + by_0 = c$. Questa é una equazione diofantea, essendo costituita da soli coefficienti interi, e la coppia (x_0, y_0) ne é soluzione. Tale equazione é infatti risolubile perché essendo $\text{MCD}(a, b)\tilde{c} = c$, si ha $c \mid \text{MCD}(a, b)$. □

Il Corollario 2.5.1 suggerisce che per ricavare una soluzione particolare di una equazione diofantea risolubile $ax + by = c$ sia sufficiente trovare una soluzione particolare dell'identitá di Bézout che ha a e b per termini noti e moltiplicare il risultato per $\frac{c}{\text{MCD}(a,b)}$.

Esempio 2.5.3: Si consideri l'equazione diofantea risolubile $74x + 22y = 10$. É giá stato calcolato che $\text{MCD}(74, 22) = 2$, pertanto l'identitá di Bézout che ha 74 e 22 come termini noti é $74x' + 22y' = 2$. Se ne determini una soluzione particolare $({x_0}', {y_0}')$:

$$\begin{aligned} 74 &= 22 \cdot 3 + 8 \Rightarrow a = 3b + 8 \Rightarrow a - 3b = 8 \\ 22 &= 8 \cdot 2 + 6 \Rightarrow b = 2(a - 3b) + 6 \Rightarrow 7b - 2a = 6 \\ 8 &= 6 \cdot 1 + 2 \Rightarrow (a - 3b) = (7b - 2a) + 2 \Rightarrow 3a - 10b = 2 \end{aligned}$$

Si ha quindi $({x_0}', {y_0}') = (3, -10)$. Essendo $\frac{10}{\text{MCD}(74,22)} = 5$, si ha che una soluzione particolare dell'equazione diofantea $74x + 22y = 10$ é $(15, -50)$.

Teorema 2.5.2 (Soluzioni di una equazione diofantea): Si consideri l'equazione diofantea risolubile $ax + by = c$, con termini noti non nulli $a, b, c \in \mathbb{Z}$ e incognite $x, y \in \mathbb{Z}$. Se la coppia $(x_0, y_0) \in \mathbb{Z} \times \mathbb{Z}$ é soluzione per tale equazione, allora lo sono tutte e sole le coppie $(x_h, y_h) \in \mathbb{Z} \times \mathbb{Z}$ cosí costruite:

$$x_h = x_0 + h\left(\frac{b}{\text{MCD}(a, b)}\right) \quad y_h = y_0 - h\left(\frac{a}{\text{MCD}(a, b)}\right) \quad \text{con } h \in \mathbb{Z}$$

Dimostrazione: Le coppie (x_h, y_h) cosí costruite sono certamente soluzioni di $ax + by = c$, dato che sostituendo si ha:

$$\begin{aligned} ax_h + by_h = c &\Rightarrow a\left(x_0 + h\left(\frac{b}{\text{MCD}(a, b)}\right)\right) + b\left(y_0 - h\left(\frac{a}{\text{MCD}(a, b)}\right)\right) = c \\ &\Rightarrow ax_0 + \cancel{\frac{ahb}{\text{MCD}(a, b)}} + by_0 - \cancel{\frac{ahb}{\text{MCD}(a, b)}} = c \Rightarrow ax_0 + by_0 = c \end{aligned}$$

Viceversa, sia $(\overline{x}, \overline{y})$ una generica soluzione di $ax + by = c$. Dato che anche (x_0, y_0) lo é, é possibile scrivere:

$$a\overline{x} + b\overline{y} = c = ax_0 + by_0 \Rightarrow a(\overline{x} - x_0) = -b(\overline{y} - y_0) \Rightarrow \overline{a}(\overline{x} - x_0) = \overline{b}(y_0 - \overline{y}) \quad \text{con} \quad \begin{matrix} \overline{a} = \frac{a}{\text{MCD}(a,b)} \\ \overline{b} = \frac{b}{\text{MCD}(a,b)} \end{matrix}$$

Dall'espressione si ricava che $\overline{a} \mid \overline{b}(y_0 - \overline{y})$, da cui si ha $\overline{a} \mid y_0 - \overline{y}$. Ma allora esiste un certo $h \in \mathbb{Z}$ tale per cui $y_0 - \overline{y} = h\overline{a}$, cioé $\overline{y} = y_0 - h\overline{a}$. Sostituendo nella precedente, si ha:

$$\overline{a}(\overline{x} - x_0) = \overline{b}(\cancel{y_0} - \cancel{y_0} + h\overline{a}) \Rightarrow \cancel{\overline{a}}(\overline{x} - x_0) = \overline{b}h\cancel{\overline{a}} \Rightarrow \overline{x} - x_0 = \overline{b}h \Rightarrow \overline{x} = x_0 + \overline{b}h$$

Risostituendo il valore di $\overline{a}$ e $\overline{b}$ nelle rispettive formule, si ottiene la forma presente nell'enunciato del teorema:

$$\overline{x} = x_0 + h\left(\frac{b}{\mathrm{MCD}(a,b)}\right) \quad \overline{y} = y_0 - h\left(\frac{a}{\mathrm{MCD}(a,b)}\right) \quad \text{con } h \in \mathbb{Z}$$

Essendo $(\overline{x}, \overline{y}) \in \mathbb{Z} \times \mathbb{Z}$ una soluzione generica, si ha quindi che qualsiasi soluzione puó essere espressa in tale forma. □

Esempio 2.5.4: Si consideri l'equazione diofantea risolubile $74x + 22y = 10$, del quale é nota la soluzione particolare $(15, -50)$ ed é noto che $\mathrm{MCD}(74, 22) = 2$. Avendosi $\frac{74}{2} = 37$ e $\frac{22}{2} = 11$, é possibile ricavare la famiglia di soluzioni $(x_h, y_h) \in \mathbb{Z} \times \mathbb{Z}$:

$$x_h = 15 + 11h \quad y_h = -50 - 37h \quad \text{con } h \in \mathbb{Z}$$

2.6. Congruenza Modulo n

Sia $n \in \mathbb{Z}$ con $n > 0$. Dati due interi a e b sono **congrui modulo** n se $n \mid a - b$, e si scrive $a \equiv b \bmod n$. In altre parole, $a \equiv b \bmod n$ vale se e solo se esiste un certo $k \in \mathbb{Z}$ tale per cui $a - b = nk$. In maniera equivalente, é possibile dire che due numeri a e b sono congruenti modulo n se la loro divisione per n restituisce il medesimo resto.

Esempio 2.6.1: Avendosi $12 \mid 38 - 14$, é possibile scrivere $38 \equiv 14 \bmod 12$. Si noti inoltre come sia 38 sia 14, divisi per 12, diano resto 2.

La definizione puó essere estesa anche al caso in cui $n = 0$. Si noti infatti come, se vale $n = 0$, si ha $a - b = 0 \cdot k$, ovvero $a = b$. Pertanto, la congruenza modulo 0 coincide semplicemente con la relazione di uguaglianza in $\mathbb{Z}$. La definizione puó essere inoltre estesa anche al caso in cui $n < 0$. Infatti, basta osservare che $n \mid a - b$ se e solo se $-n \mid a - b$ per concludere che $a \equiv b \bmod n$ se e solo se $a \equiv b \bmod -n$. Per questo motivo, non é limitativo considerare $n > 0$.

Lemma 2.6.1: Sia $n \in \mathbb{Z}$ con $n > 0$. Dati quattro interi a, b, c e d, se vale $a \equiv b \bmod n$ e $c \equiv d \bmod n$ allora vale $a + b \equiv c + d \bmod n$.

Lemma 2.6.2: Sia $n \in \mathbb{Z}$ con $n > 0$. Dati quattro interi a, b, c e d, se vale $a \equiv b \bmod n$ e $c \equiv d \bmod n$ allora vale $ab \equiv cd \bmod n$.

Esempio 2.6.2: La congruenza lineare dell'Esempio 2.7.3, che aveva per soluzione particolare $c = 5$. Avendosi $\mathrm{MCD}(21, 30) = 3$, si ha $\frac{30}{3} = 10$. Pertanto, tale congruenza lineare ha per soluzioni ogni intero nella forma $6 + 10h$ con $h \in \mathbb{Z}$. In particolare, le soluzioni non congruenti modulo n fra di loro sono $c = 6$, $c = 16$ e $c = 26$.

Lemma 2.6.3: Siano $a, b, c, n \in \mathbb{Z}$, con $c \neq 0$. Allora $ac \equiv bc \bmod n$ equivale a $a \equiv b \bmod \frac{n}{\mathrm{MCD}(c,n)}$.

Dimostrazione: Per definizione di congruenza modulo n, l'espressione $ac \equiv bc \bmod n$ equivale a $n \mid ac - bc$. Deve allora esistere un certo $q \in \mathbb{Z}$ tale per cui $ac - bc = nq$, ovvero $(a-b)c = nq$. Siano $c = \tilde{c}\ \mathrm{MCD}(c,n)$ e $n = \tilde{n}\ \mathrm{MCD}(c,n)$. Si ha:

$$(a-b)c = nq \Rightarrow (a-b)\tilde{c}\cancel{\mathrm{MCD}(c,n)} = \tilde{n}\cancel{\mathrm{MCD}(c,n)}q \Rightarrow (a-b)\tilde{c} = \tilde{n}q \Rightarrow \tilde{n} \mid (a-b)\tilde{c}$$

Per il Lemma 2.4.1, almeno una delle due proposizioni fra $\tilde{n} \mid a-b$ e $\tilde{n} \mid \tilde{c}$ deve essere vera. La prima proposizione equivale a $a \equiv b \bmod \tilde{n}$; ricordando la definizione di $\tilde{n}$, si ha $a \equiv b \bmod \frac{n}{\mathrm{MCD}(c,n)}$. □

Corollario 2.6.1 (Legge di cancellazione per le congruenze lineari): Siano $a, b, c, n \in \mathbb{Z}$, con c non nullo e con c ed n coprimi. Allora $ac \equiv bc \bmod n$ equivale a $a \equiv b \bmod n$.

Dimostrazione: Se c ed n sono coprimi, allora $\mathrm{MCD}(c,n) = 1$. Applicando il Lemma 2.6.3, si ha che $ac \equiv bc \bmod n$ equivale a $a \equiv b \bmod \frac{n}{1}$, ovvero $a \equiv b \bmod n$. □

Teorema 2.6.1: Per ogni numero intero $n > 0$, la congruenza modulo n é una relazione di equivalenza su $\mathbb{Z}$.

Dimostrazione: La congruenza modulo n definisce su $\mathbb{Z}$ la relazione $\mathcal{R}$ data da:

$$\forall a, b \in \mathbb{Z}, (a,b) \in \mathcal{R} \text{ se e solo se } a \equiv b \bmod n$$

La relazione in questione é:

1. Riflessiva: $\forall a \in \mathbb{Z}$ vale $a \equiv a \bmod n$. Infatti, $a \equiv a \bmod n$ equivale a dire $a - a = 0 = kn$, che é valido per $k = 0$ e per qualsiasi $a \in \mathbb{Z}$;
2. Simmetrica: $\forall a, b \in \mathbb{Z}$, $a \equiv b \bmod n$ implica $b \equiv a \bmod n$. Infatti, $a \equiv b \bmod n$ equivale a dire $a - b = kn$ per un certo $k \in \mathbb{Z}$. Moltiplicando per -1 ambo i membri si ha $-(a-b) = -(kn)$, ovvero $b - a = (-k)n$, cioé $b \equiv a \bmod n$;
3. Transitiva: $\forall a, b, c \in \mathbb{Z}$, $a \equiv b \bmod n$ e $b \equiv c \bmod n$ implicano $a \equiv c \bmod n$. Infatti, $a \equiv b \bmod n$ e $b \equiv c \bmod n$ equivalgono a dire, rispettivamente, $a - b = kn$ e $b - c = hn$ per certi $h, k \in \mathbb{Z}$. Sommando la seconda alla prima:

$$a - b + (b - c) = kn + (b - c) \Rightarrow a - \cancel{b} + \cancel{b} - c = kn + hn \Rightarrow a - c = (k+h)n \Rightarrow a \equiv c \bmod n$$

Pertanto, é una relazione di equivalenza. □

Essendo la congruenza modulo n una relazione di equivalenza, é possibile identificare delle classi di equivalenza. Preso n intero con $n > 0$ ed un certo $a \in \mathbb{Z}$, la classe di equivalenza di a rispetto alla congruenza modulo n viene indicata con $[a]_n$.
Tale classe di equivalenza corrisponde all'insieme $\{b : b \in \mathbb{Z} \wedge a \equiv b \bmod n\}$, ovvero all'insieme che contiene tutti i numeri interi che, divisi per n, restituiscono lo stesso resto della divisione fra n e a.

Lemma 2.6.4: Sia n un numero intero maggiore di 0. Sia a un numero intero qualsiasi e sia b il resto della divisione di a per n. Vale $[a]_n = [b]_n$.

Dimostrazione: Se b é il resto della divisione di a per n, allora vale $a = nk + b$ per un certo $k \in \mathbb{Z}$, da cui si ha $a - b = nk$, che é la definizione di congruenza modulo n. □

L'insieme quoziente di $\mathbb{Z}$ rispetto alla relazione di congruenza modulo n con $n > 0$ si dice **insieme delle classi di resti modulo** n e si denota con $\mathbb{Z}_n$.

Teorema 2.6.2: Per ogni numero intero $n > 0$, l'insieme delle classi di resti modulo n ha cardinalitá n. In particolare, tale insieme é:

$$\mathbb{Z}_n = \left\{[0]_n, [1]_n, ..., [n-1]_n\right\} = \{\{nk : k \in \mathbb{Z}\}, \{1 + nk : k \in \mathbb{Z}\}, ..., \{n - 1 + nk : k \in \mathbb{Z}\}\}$$

Dimostrazione: Sia $a \in \mathbb{Z}$. La divisione con resto fornisce $a = nq + r$ con $0 \leq r < n$. Poichè $a - r = qn$ si ha che $a \equiv r \bmod n$. Ciò mostra che ogni intero a è congruo, modulo n, a uno degli interi $0, 1, ..., n-1$. D'altra parte se i e j sono interi, con $0 \leq i < n$ e $0 \leq j < n$ si ha, assumendo $i \geq j$, che $0 \leq i - j \leq n - 1$ e quindi $i - j = kn$ se e solo se $k = 0$, cioè $i = j$. □

Esempio 2.6.3:

Con $n = 2$, si ha $\mathbb{Z}_2 = \left\{[0]_2, [1]_2\right\}$:

$$[0]_2 = \{..., -6, -4, -2, 0, 2, 4, 6, ...\}$$
$$[1]_2 = \{..., -5, -3, -1, 1, 3, 5, 7, ...\}$$

Con $n = 3$, si ha $\mathbb{Z}_3 = \left\{[0]_3, [1]_3, [2]_3\right\}$:

$$[0]_3 = \{..., -9, -6, -3, 0, 3, 6, 9, ...\}$$
$$[1]_3 = \{..., -8, -5, -2, 1, 4, 7, 10, ...\}$$
$$[2]_3 = \{..., -7, -4, -1, 2, 5, 8, 11, ...\}$$

Ad esempio, la classe di resto $[5]_7$ rappresenta, oltre al numero 5, anche il numero 12 $(1 \times 7 + 5)$, il numero 19 $(2 \times 7 + 5)$, il numero 2308 $(329 \times 7 + 5)$, il numero -2 $(-1 \times 7 + 5)$ il numero -9 $(-2 \times 7 + 5)$, ecc...

Lemma 2.6.5: Sia $[a]_n$ con $n \in \mathbb{N}$ una classe di resto. Se vale $[a]_n = [0]_n$, allora $n \mid a$.

Dimostrazione: Per la definizione di classe di resto, l'espressione $[a]_n = [0]_n$ equivale a dire che la divisione fra a e n ha lo stesso resto della divisione fra 0 ed n. Dato che la divisione fra 0 ed un qualsiasi numero (intero) ha quoziente 0 e resto 0, si ha che la divisione fra a e n ha resto 0, ovvero che $n \mid a$. □

Sull'insieme delle classi di resto modulo n é possibile definire le operazioni di somma e di prodotto. Siano $[a]_n$ e $[b]_n$ due classi di resto modulo n. La somma ed il prodotto sono definiti come:

$$[a]_n + [b]_n = [a + b]_n \quad [a]_n \cdot [b]_n = [ab]_n$$

Esempio 2.6.4: in $\mathbb{Z}_5$, si ha $[1]_5 + [3]_5 = [3 + 1]_5 = [4]_5$ e $[2]_5 \cdot [3]_5 = [2 \cdot 3]_5 = [6]_5$

Lemma 2.6.6: Sia $n \in \mathbb{Z}$ con $n \geq 1$. Siano poi $a, b, c, d \in \mathbb{Z}$, tali per cui $[a]_n = [b]_n$ e $[c]_n = [d]_n$. Allora vale:

$$[a]_n + [c]_n = [b]_n + [d]_n \quad [a]_n \cdot [c]_n = [b]_n \cdot [d]_n$$

Dimostrazione: Poichè $[a]_n = [b]_n$ e $[c]_n = [d]_n$ si ha, per definizione di classe di equivalenza, $a = b + nk$ e $c = d + nh$ per $k, h \in \mathbb{Z}$. Sommando e moltiplicando l'una all'altra, si ha:

$$a + c = b + nk + d + nh \Rightarrow a + c = (b + d) + n(h + k)$$
$$a \cdot c = (b + nk) \cdot (d + nh) \Rightarrow ac = bd + bnh + nkd + n^2kh \Rightarrow ac = bd + n(bh + dk + nkh)$$

Essendo $\mathbb{Z}$ chiuso rispetto alla somma e al prodotto, si ha $k + h \in \mathbb{Z}$ e $bh + dk + khn \in \mathbb{Z}$, siano questi rispettivamente α e β. Si ha:

$$(a + c) = (b + d) + n\alpha \quad ac = bd + n\beta$$

Applicando nuovamente la definizione di classe di equivalenza, si ha che $[a + c]_n = [b + d]_n$ e $[ac]_n = [bd]_n$. Per come sono state definite la somma ed il prodotto rispetto alle classi di equivalenza, si ha infine $[a]_n + [c]_n = [b]_n + [d]_n$ e $[a]_n[c]_n = [b]_n[d]_n$. □

Teorema 2.6.3: La struttura algebrica $(\mathbb{Z}_n, +)$, formata dalla classe di resti modulo n e dalla somma su questi definita, é un gruppo abeliano.

Dimostrazione: La struttura algebrica $(\mathbb{Z}_n, +)$ é:

- un semigruppo, perché l'operazione $+$ cosí definita gode della proprietá associativa. Questo é determinato dal fatto che l'usuale somma in $\mathbb{Z}$ gode di tale proprietá:

$$\begin{aligned}\left([a]_n + [b]_n\right) + [c]_n = [a + b]_n + [c]_n = [(a + b) + c]_n = [a + (b + c)]_n \\ = [a]_n + [b + c]_n = [a]_n + \left([b]_n + [c]_n\right)\end{aligned}$$

- un monoide, perché per l'operazione $+$ cosí definita esiste l'elemento neutro. Tale elemento é $[0]_n$, infatti preso un qualsiasi $[a]_n \in \mathbb{Z}_n$:

$$[0]_n + [a]_n = [a]_n + [0]_n = [a + 0]_n = [0 + a]_n = [a]_n$$

- un gruppo, perché per l'operazione $+$ cosí definita esiste un elemento opposto per qualsiasi elemento di $\mathbb{Z}_n$. Preso un qualsiasi $[a]_n \in \mathbb{Z}_n$, tale elemento opposto é $[n - a]_n$, in quanto:

$$[a]_n + [n - a]_n = [n - a]_n + [a]_n = [(n - a) + a]_n = [a + (n - a)]_n = [n]_n = 0$$

Inoltre, $+$ gode della proprietá commutativa. Infatti:

$$[a]_n + [b]_n = [a + b]_n = [b + a]_n = [b]_n + [a]_n$$

Pertanto, $(\mathbb{Z}_n, +)$ é un gruppo abeliano. □

Teorema 2.6.4: La struttura algebrica $(\mathbb{Z}_n, \cdot)$, formata dalla classe di resti modulo n e dal prodotto su questi definito, é un monoide abeliano.

Dimostrazione: La struttura algebrica $(\mathbb{Z}_n, \cdot)$ é:

- un semigruppo, perché l'operazione $\cdot$ cosí definita gode della proprietá associativa. Questo é determinato dal fatto che l'usuale prodotto in $\mathbb{Z}$ gode di tale proprietá:

$$\begin{aligned}\left([a]_n \cdot [b]_n\right) \cdot [c]_n = [a \cdot b]_n \cdot [c]_n = [(a \cdot b) \cdot c]_n = [a \cdot (b \cdot c)]_n \\ = [a]_n \cdot [b \cdot c]_n = [a]_n \cdot \left([b]_n \cdot [c]_n\right)\end{aligned}$$

- un monoide, perché per l'operazione $\cdot$ cosí definita esiste l'elemento neutro. Tale elemento é $[1]_n$, infatti preso un qualsiasi $[a]_n \in \mathbb{Z}_n$:

$$[1]_n \cdot [a]_n = [a]_n \cdot [1]_n = [a \cdot 1]_n = [1 \cdot a]_n = [a]_n$$

Inoltre, $\cdot$ gode della proprietá commutativa. Infatti:

$$[a]_n \cdot [b]_n = [a \cdot b]_n = [b \cdot a]_n = [b]_n \cdot [a]_n$$

Pertanto, $(\mathbb{Z}_n, \cdot)$ é un monoide abeliano. □

Un elemento $[a]_n$ in $\mathbb{Z}_n$ si dice **invertibile** in $\mathbb{Z}_n$ (rispetto al prodotto) se esiste $[b]_n \in \mathbb{Z}_n$ tale per cui $[a]_n \cdot [b]_n = [1]_n$; $[b]_n$ viene detto *inverso* di $[a]_n$.

Esempio 2.6.5:
- Se vale $[a]_7 = [3]_7$, allora esiste $[b]_7 \in \mathbb{Z}_7$ tale per cui $[3]_7 \cdot [b]_7 = [1]_7$. Tale $[b]_7$ é $[5]_7$, in quanto $[3]_7[5]_7 = [15]_7 = [1]_7$;
- Se vale $[a]_6 = [3]_6$, allora non esiste alcun $[b]_6 \in \mathbb{Z}_6$ tale per cui $[3]_6[b]_6 = [1]_6$;
- Se vale $[a]_n = [0]_n$ per un qualsiasi $n \in \mathbb{N}$, per ogni $[b]_n \in \mathbb{Z}_n$, risulta $[0]_n[b]_n = [0 \cdot b]_n = [0]_n$. Pertanto, affinché esista un $[b]_n \in \mathbb{Z}_n$ tale per cui valga $[0]_n[b]_n = [1]_n$, deve valere $[0]_n = [1]_n$, ovvero $0 \equiv 1 \bmod n$. Questo si verifica soltanto se $n = 1$, pertanto un invertibile per $[a]_n = [0]_n$ esiste solamente in questo caso.

Lemma 2.6.7: Siano a, n due numeri interi, dove $n \geq 2$. La classe di resto $[a]_n$ è invertibile in $\mathbb{Z}_n$ se e soltanto se a ed n sono coprimi, ovvero se $\text{MCD}(a, n) = 1$.

Dimostrazione: Se la classe di resto $[a]_n$ è invertibile, allora esiste $[b]_n \in \mathbb{Z}_n$ tale per cui $[a]_n[b]_n = [1]_n$, ovvero $[ab]_n = [1]_n$. Per come la somma sulle classi di resto é stata definita, é possibile sommare $[-1]_n$ ad entrambi i membri, ottenendo $[ab]_n + [-1]_n = [1]_n + [-1]_n$, da cui si ricava $[ab - 1]_n = [0]_n$. Per il Lemma 2.6.5, si ha $n \mid ab - 1$. Deve allora esistere un $k \in \mathbb{Z}$ tale per cui $ab - 1 = nk$, ovvero $ab - nk = 1$. Dato che sia b sia k sono certamente esistenti, é possibile applicare il Lemma 2.2.1 per provare che a ed n sono coprimi.
Viceversa, si assuma che a ed n siano coprimi. Per l'identitá di Bézout esistono $s, t \in \mathbb{Z}$ tali per cui $as + nt = 1$, ovvero $as = 1 - nt$. Questo equivale a dire che $as \equiv 1 \bmod n$, ovvero che $[as]_n = [a]_n[s]_n = [1]_n$. Si ha quindi che per $[a]_n$ esiste l'invertibile. □

Esempio 2.6.6: In $\mathbb{Z}_{51}$ l'elemento $[13]_{51}$ è invertibile perchè $\text{MCD}(13, 51) = 1$. D'altro canto, $[15]_{51}$ non lo é, perché $\text{MCD}(15, 51) = 3$.

Lemma 2.6.8: Se la classe di resto $[a]_n$ é invertibile, il suo inverso é unico.

L'inverso di una classe di resto $[a]_n$, essendo unico, viene anche indicato semplicemente con $[a]_n^{-1}$.

Lemma 2.6.9: Sia $\mathbb{Z}_n$ un insieme di classi di resto modulo n, con n numero primo. Tutte le classi di resto di $\mathbb{Z}_n$, tranne $[0]_n$, sono invertibili.

2.7. Congruenze lineari

Viene detta **congruenza lineare modulo** n qualunque espressione nella forma:

$$ax \equiv b \bmod n \qquad \text{con } a, b, n \in \mathbb{Z}$$

Dove a, b ed n sono termini noti ed x é una incognita. Naturalmente, le *soluzioni* di una congruenza lineare sono tutti e soli quei $c \in \mathbb{Z}$ tali che, sostituiti ad x, rendono valida l'espressione. Se esiste almeno un c con queste caratteristiche, si dice che la congruenza lineare *ammette* soluzione.

Esempio 2.7.1: Si consideri la congruenza lineare $2x \equiv 3 \bmod 7$. Una possibile soluzione per tale congruenza é $c = 5$, dato che $2 \cdot 5 = 10$ ed effettivamente $10 \equiv 3 \bmod 7$. Anche $c = 26$ é una possibile soluzione, dato che $2 \cdot 26 = 52 \equiv 3 \bmod 7$.

Teorema 2.7.1: Siano $a, b, n \in \mathbb{Z}$, con $a \neq 0$. La congruenza lineare $ax \equiv b \bmod n$ ammette soluzione se e soltanto se $\text{MCD}(a, n) \mid b$.

Dimostrazione: Da definizione di congruenza modulo n, si ha che $ax \equiv b \bmod n$ equivale a $n \mid ax - b$, che a sua volta equivale a $ax - b = nk$ per un certo $k \in \mathbb{Z}$. Spostando b al secondo membro, si ha $ax - nk = b$; dato che tutti i numeri che figurano in questa equazione sono numeri interi, si sta avendo a che fare con una equazione diofantea, nello specifico nelle variabili x e k. Per il Teorema 2.5.1, l'equazione ha soluzione se e soltanto se $\text{MCD}(a, n) \mid b$, ma dato che tale equazione é solamente una riscrittura di $ax \equiv b \bmod n$, allora anche quest'ultima avrá soluzione se e solo se sono rispettate tali condizioni. □

Esempio 2.7.2:
- La congruenza lineare dell'Esempio 2.7.1 ha soluzioni, perché $\text{MCD}(a, n) = 1$ ed é vero che $1 \mid 3$;
- La congruenza lineare $2x \equiv 3 \bmod 4$ non ha soluzioni, perché $\text{MCD}(a, n) = 2$ ed é falso che $2 \mid 3$.

Il Teorema 2.7.1 fornisce implicitamente un approccio per cercare una soluzione particolare di una congruenza lineare, ovvero costruendo una equazione diofantea a questa equivalente e risolvendola. La soluzione particolare é data dalla componente x della soluzione particolare di tale equazione.

Esempio 2.7.3: Si consideri la congruenza lineare $21x \equiv 6 \bmod 30$. L'equazione diofantea associata é $21x - 30k = 6$. Si ha:

$$\begin{array}{l} 30 = 21 \cdot 1 + 9 \\ 21 = 9 \cdot 2 + 3 \\ 9 = 3 \cdot 3 + 0 \end{array} \quad \begin{array}{l} b = a \cdot 1 + 9 \Rightarrow 9 = b - a \\ a = 2(b - a) + 3 \Rightarrow 3 = 3a - 2b \end{array} \quad (6)21 - (4)30 = 6$$

Da cui si ricava la soluzione particolare $c = 6$ per la congruenza lineare.

Teorema 2.7.2: Siano $a, b, n \in \mathbb{Z}$, con $a \neq 0$. Si consideri la congruenza lineare $ax \equiv b \bmod n$: se $x_0 \in \mathbb{Z}$ ne é una soluzione, allora lo sono anche tutti ed i soli numeri interi x_h nella forma:

$$x_h = x_0 + h\left(\frac{n}{\text{MCD}(a, n)}\right) \quad \text{con } h \in \mathbb{Z}$$

In particolare, fra queste ne esistono esattamente $\text{MCD}(a, n)$ non congruenti modulo n fra di loro.

Dimostrazione: Per il Teorema 2.7.1, $ax \equiv b \bmod n$ ha soluzione se e soltanto se ha soluzione l'equazione diofantea equivalente $ax - nk = b$ con $k \in \mathbb{Z}$. Per il Teorema 2.5.2 si ha che se $(x_0, k_0) \in \mathbb{Z} \times \mathbb{Z}$ é una soluzione particolare di tale equazione, allora lo sono tutte e sole le coppie $(x_h, k_h) \in \mathbb{Z} \times \mathbb{Z}$ nella forma:

$$x_h = x_0 + h\left(\frac{n}{\mathrm{MCD}(a,n)}\right) \quad k_h = k_0 - h\left(\frac{n}{\mathrm{MCD}(a,n)}\right) \quad \text{con } h \in \mathbb{Z}$$

L'espressione per x_h é quella cercata. Per provare che la congruenza lineare ha esattamente $\mathrm{MCD}(a,n)$ soluzioni non congruenti modulo n fra di loro, si consideri $h_1, h_2 \in \mathbb{Z}$. Si ha:

$$x_0 + h_1\left(\frac{n}{\mathrm{MCD}(a,n)}\right) \equiv x_0 + h_2\left(\frac{n}{\mathrm{MCD}(a,n)}\right) \bmod n \iff \left(\frac{n}{\mathrm{MCD}(a,n)}\right)(h_1 - h_2) \equiv 0 \bmod n$$

Deve allora esistere un certo $q \in \mathbb{Z}$ tale per cui:

$$\left(\frac{n}{\mathrm{MCD}(a,n)}\right)(h_1 - h_2) \equiv 0 \bmod n \Rightarrow \left(\frac{\cancel{n}}{\mathrm{MCD}(a,n)}\right)(h_1 - h_2) = q\cancel{n} \Rightarrow h_1 - h_2 = q\ \mathrm{MCD}(a,n)$$

Pertanto, le $\mathrm{MCD}(a,n)$ soluzioni non congruenti modulo n fra di loro che si stavano cercando sono tutte e sole le soluzioni con $h = 0, 1, ..., (\mathrm{MCD}(a,n) - 1)$. □

Sia $[a]_n$ una classe di resto invertibile, e si supponga di volerne trovarne l'inverso $[a]_n^{-1}$. É sufficiente osservare come l'espressione $[a]_n[a]_n^{-1} = [1]_n$ equivalga a $a \cdot a^{-1} \equiv 1 \bmod n$. Pertanto, occorre risolvere tale congruenza lineare con a^{-1} come incognita e sceglierne una soluzione qualsiasi, essendo tutte equivalenti. Per convenzione, si preferisce scegliere la soluzione piú piccola.

Esempio 2.7.4: In $\mathbb{Z}_9$, la classe di resto $[7]_9$ é invertibile, in quanto $\mathrm{MCD}(7,9) = 1$. L'inverso é ricavato dal risolvere la congruenza lineare $7x \equiv 1 \bmod 9$, che ha come soluzione $4 + 9k$ con $k \in \mathbb{Z}$. Pertanto, l'inverso di $[7]_9$ é $[4]_9$.

Viene detto **sistema di congruenze lineari** qualunque espressione nella forma:

$$A_i x \equiv B_i \bmod N_i = \begin{cases} a_1 x \equiv b_1 \mod n_1 \\ a_2 x \equiv b_2 \mod n_2 \\ \vdots \\ a_m x \equiv b_m \bmod n_m \end{cases} \quad \text{con } a_1, ..., a_m, b_1, ..., b_m, n_1, ..., n_m \in \mathbb{Z}$$

Dove $a_1, ..., a_m$, $b_1, ..., b_m$ e $n_1, ..., n_m$ sono termini noti ed x é una incognita. Le *soluzioni* di un sistema di congruenze lineari sono tutti e soli quei $c \in \mathbb{Z}$ tali che, sostituiti ad x, verificano contemporaneamente tutte le m congruenze lineari modulo n_i che lo compongono. Se esiste almeno un c con queste caratteristiche, si dice che il sistema di congruenze lineari *ammette* soluzione.

Lemma 2.7.1 (Condizione necessaria per la solubilitá di un sistema di congruenze lineari): Un sistema di congruenze lineari $A_i x \equiv B_i \bmod N_i$ ha soluzione soltanto se, per ogni $i = 1, ..., m$, si ha $\mathrm{MCD}(a_i, n_i) \mid b_i$.

Dimostrazione: Per il Teorema 2.7.1, si ha che $ax \equiv b \bmod n$ ha soluzione se e soltanto se $\mathrm{MCD}(a,n) \mid b$. Dato che un sistema di congruenze lineari ha soluzione soltanto se tutte le congruenze che lo compongono hanno soluzione, tale sistema avrá soluzione soltanto se $\mathrm{MCD}(a_i, n_i) \mid b_i$ é valido per ogni $i = 1, ..., m$.□

Si noti come il Lemma 2.7.1 sia una implicazione a senso unico, ovvero potrebbero esistere dei sistemi di congruenze lineari che lo verificano ma che comunque non hanno soluzione. Infatti, le congruenze lineari che costituiscono un sistema potrebbero essere solubili individualmente, ma nessuna di queste avere una soluzione che sia comune a tutte.

Teorema 2.7.3 (Teorema Cinese del Resto): Si consideri un sistema di congruenze lineari come quello presentato di seguito:

$$\begin{cases} x \equiv b_1 \mod n_1 \\ \vdots \\ x \equiv b_2 \mod n_2 \\ x \equiv b_m \mod n_m \end{cases} \quad \text{con } b_1, ..., b_m, n_1, ..., n_m \in \mathbb{Z}$$

Ovvero, dove i termini $a_1, ..., a_m$ sono tutti pari ad 1. Si assuma inoltre che i termini $n_1, ..., n_m$ siano tutti positivi e che siano a due a due coprimi, ovvero $\text{MCD}(n_i, n_j) = 1$ per ogni $1 \leq i \leq m$ e $1 \leq j \leq m$ tali per cui $i \neq j$.
Allora il sistema é risolubile. In particolare, se c e c' sono due soluzioni, allora vale:

$$c \equiv c' \bmod N \quad \text{dove } N = n_1 \cdot n_2 \cdot ... \cdot n_m = \prod_{i=1}^{m} n_i$$

Dimostrazione: Per ogni $i = 1, ..., m$, sia $N_i = \frac{N}{n_i}$ (essendo $N = \prod_{i=1}^{m} n_i$ é garantito che N_i sia un numero intero, perché n_i é uno dei fattori di N). Per ipotesi, si ha $\text{MCD}(n_i, n_j) = 1$ per $i \neq j$. Tuttavia, é facile verificare che anche $\text{MCD}(N_i, n_i) = 1$.
Infatti, si supponga per assurdo che $\text{MCD}(N_i, n_i) \neq 1$. Deve allora esistere un numero primo p tale per cui $p \mid n_i$ e $p \mid N_i$, ovvero che é divisore sia di n_i che di N_i. Essendo $N_i = n_1 \cdot ... \cdot n_{i-1} \cdot n_{i+1} \cdot ... \cdot n_m$, per il Lemma 2.4.1 deve esistere un n_j con $j \neq i$ tale per cui $p \mid n_j$. Ma allora, valendo sia $p \mid n_i$ sia $p \mid n_j$, si ha che n_i ed n_j hanno un divisore in comune, e quindi non sono primi, contro l'ipotesi che invece lo siano. Occorre allora assumere che $\text{MCD}(N_i, n_i) = 1$.
Si consideri la congruenza lineare $N_i y \equiv 1 \bmod n_i$ nell'incognita y, che ha y_i per soluzione. Per il Teorema 2.7.1, tale congruenza lineare ha soluzione se vale $\text{MCD}(N_i, n_i) \mid 1$, ed é stato appena mostrato che $\text{MCD}(N_i, n_i) = 1$, pertanto é garantito che y_i esista. Sia c definito come:

$$c = \sum_{i=1}^{m} N_i y_i b_i = N_1 y_1 b_1 + ... + N_m y_m b_m$$

É possibile verificare che c é una soluzione del sistema, ovvero che $c \equiv b_j \bmod n_j$ per $j \neq i$. Valendo $n_j \mid N_i$ per qualsiasi $j \neq i$, é possibile scrivere $N_i \equiv 0 \bmod n_j$, e quindi $c \equiv N_j y_j b_j \bmod n_j$. Avendo trovato che vale $N_j n_j \equiv 1 \bmod n_j$, moltiplicando ambo i membri per b_j si ha $N_j n_j b_j \equiv b_j \bmod n_j$ (questo é legittimo perché $N_j n_j$ e 1 sono primi fra di loro, esiste un lemma che lo prova).
Avendosi la soluzione c, sia c' un'altra soluzione del sistema. Allora deve valere $c \equiv c' \bmod n_i$, ovvero $n_i \mid c - c'$ per ogni i = 1, ..., m. Poichè gli n_i sono a due a due coprimi, segue che anche N é divisore di $c - c'$, ovvero $c \equiv c' \bmod N$. Questo dimostra che c è l'unica soluzione del sistema modulo N, a meno di multipli di N. □

Esempio 2.7.5: Si consideri il seguente sistema di congruenze lineari, e lo si risolva:

$$\begin{cases} x \equiv 2 \bmod 3 \\ x \equiv 3 \bmod 5 \\ x \equiv 2 \bmod 7 \end{cases}$$

Tale sistema rispetta le ipotesi del Teorema 2.7.3, dato che tutti i termini noti a sinistra dell'equivalenza sono pari ad 1, i termini noti a destra sono tutti positivi e sono tutti coprimi fra di loro a due a due.
Si ha allora $N = 3 \cdot 5 \cdot 7 = 105$. Per ciascuna congruenza lineare del sistema si calcoli $N_i = \frac{N}{n_i}$:

$$N_1 = \frac{N}{n_1} = \frac{105}{3} = 35 \quad N_2 = \frac{N}{n_2} = \frac{105}{5} = 21 \quad N_3 = \frac{N}{n_3} = \frac{105}{7} = 15$$

Da cui si ottengono le congruenze lineari:

$$\begin{aligned} N_1 y \equiv 1 \bmod n_1 &\Rightarrow 35y \equiv 1 \bmod 3 \Rightarrow 2y \equiv 1 \bmod 3 \Rightarrow y_1 = 2 \\ N_2 y \equiv 1 \bmod n_2 &\Rightarrow 21y \equiv 1 \bmod 5 \Rightarrow y \equiv 1 \bmod 5 \Rightarrow y_2 = 1 \\ N_3 y \equiv 1 \bmod n_3 &\Rightarrow 15y \equiv 1 \bmod 7 \Rightarrow y \equiv 1 \bmod 7 \Rightarrow y_3 = 1 \end{aligned}$$

La soluzione del sistema é allora data da:

$$c = \sum_{i=1}^{3} N_i y_i b_i = 35 \cdot 2 \cdot 2 + 21 \cdot 1 \cdot 3 + 15 \cdot 1 \cdot 2 = 233$$

E da tutti gli interi a questo congruenti modulo 105.

2.8. Funzione di Eulero

Viene detta **funzione di Eulero** la funzione $\varphi : (\mathbb{N} - \{0\}) \mapsto (\mathbb{N} - \{0\})$ cosí definita:

$$\varphi(n) = \begin{cases} 1 & \text{se } n = 1 \\ |\{k \in \mathbb{N} : 0 < k < n, \text{MCD}(k, n) = 1\}| & \text{se } n > 1 \end{cases}$$

Ovvero, che per l'argomento 1 restituisce 1 mentre per un generico argomento n, numero naturale maggiore di 1, restituisce il numero di numeri naturali coprimi ad n che si trovano nell'intervallo $(0, n)$, estremi esclusi.

Esempio 2.8.1: Per $n = 26$, si ha:

$$\varphi(26) = |\{k \in \mathbb{Z} : 0 < k < 26, (k, 26) = 1\}| = |\{1, 3, 5, 7, 9, 11, 15, 17, 19, 21, 23, 25\}| = 12$$

Lemma 2.8.1: Se $p \in \mathbb{N}$ é un numero primo maggiore di 1, allora $\varphi(p) = p - 1$.

Dimostrazione: Per un generico p numero naturale con $p > 1$, $\varphi(p)$ é il numero di numeri naturali maggiori di 0 e minori di p con cui p é coprimo. Se peró p é primo, allora sará certamente coprimo a tutti i numeri che costituiscono tale intervallo; essendo tale intervallo di lunghezza $p - 1$, si ha $\varphi(p) = p - 1$. □

Lemma 2.8.2: Siano p e α due numeri naturali maggiori di 0, con p primo. Allora:

$$\varphi(p^\alpha) = p^\alpha - p^{\alpha-1} = p^{\alpha-1}(p - 1)$$

Dimostrazione: Sia m un qualsiasi numero naturale diverso da 0 e inferiore a p^α. Essendo p un numero primo, gli unici possibili valori di $\text{MCD}(p^\alpha, m)$ sono $p^0, p^1, p^2, ..., p^{\alpha-1}$. Affinché $\text{MCD}(p^\alpha, m)$ non sia 1,

m deve necessariamente essere un multiplo di p, ed il numero di multipli p minori di p^α é $p^{\alpha-1}$. Tutti i restanti numeri compresi (estremi esclusi) fra 0 e p^α sono coprimi a p^α, ed il numero di tali numeri deve quindi essere $p^\alpha - p^{\alpha-1}$. □

Teorema 2.8.1 (Moltiplicativitá della funzione di Eulero): La funzione di Eulero é moltiplicativa. Ovvero, presi $a, b \in \mathbb{N} - \{0\}$ primi fra di loro, si ha $\varphi(ab) = \varphi(a)\varphi(b)$.

Dimostrazione: Siano r e s due numeri interi, scelti con queste caratteristiche:

$$0 < r < a \quad \text{MCD}(r, a) = 1 \quad 0 < s < b \quad \text{MCD}(s, b) = 1$$

Per il Teorema 2.7.3, il sistema di congruenze

$$\begin{cases} x \equiv r \bmod a \\ x \equiv s \bmod b \end{cases}$$

ammette soluzioni. In particolare, ne ammette una ed una sola compresa tra 0 e ab (estremi esclusi); sia c questa soluzione.
É possibile verificare che $\text{MCD}(c, ab) = 1$. Si assuma infatti per assurdo che questo non sia vero, e che esista pertanto un numero primo p divisore sia di c che di ab. Valendo $p \mid ab$, é possibile applicare il Lemma 2.4.1, pertanto deve valere almeno un assunto fra $p \mid a$ e $p \mid b$. Si supponga che sia vera $p \mid a$. Essendo c soluzione del sistema di congruenze, deve valere $c \equiv r \bmod a$, ovvero che esiste un $k \in \mathbb{Z}$ tale per cui $c - r = ak$. Riscrivendo l'espressione come $r = c - ah$, si evince che $p \mid r$, ma si ha assunto che valesse $p \mid a$ e che $\text{MCD}(r, a) = 1$, e le due assunzioni sono incompatibili. É facile verificare che assumendo invece che sia vera $p \mid b$, si ricade in una contraddizione analoga, pertanto occorre assumere che effettivamente $\text{MCD}(c, ab) = 1$.
Poiché ogni coppia di interi r ed s definiti come sopra dá luogo ad un intero c tale che $0 < c < ab$ e $\text{MCD}(c, ab) = 1$ abbiamo che $\varphi(a)\varphi(b) \leq \varphi(ab)$.
Viceversa, sia t un numero intero scelto di modo che valga $0 < t < ab$ e $\text{MCD}(t, ab) = 1$. Dividendo t per a, si ha $t = aq + r$ con $0 \leq r < a$ e $q \in \mathbb{Z}$.
É possibile verificare che $\text{MCD}(a, r) = 1$. Innanzitutto, si osservi come debba per forza aversi $r \neq 0$; se cosí fosse, si avrebbe $a \mid t$, ma questo non é possibile perché per come t é stato definito deve valere $\text{MCD}(t, ab) = 1$. Si supponga per assurdo che $\text{MCD}(a, r) > 1$: se cosí fosse, deve valere sia $\text{MCD}(a, r) \mid a$ che $\text{MCD}(a, r) \mid r$, da cui si ha $\text{MCD}(a, r) \mid ab$ e $\text{MCD}(a, r) \mid t$, che é una contraddizione. Occorre pertanto assumedere che effettivamente $\text{MCD}(a, r) = 1$.
In maniera analoga, si mostra che dividendo t per b e scrivendo $t = b\overline{q} + s$ con $0 < s \leq b$ si ha $\text{MCD}(b, s) = 1$. In totale, si ha che t é soluzione del sistema di congruenze:

$$\begin{cases} x \equiv r \bmod a \\ x \equiv s \bmod b \end{cases}$$

Da cui si conclude che $\varphi(a)\varphi(b) = \varphi(ab)$. □

Corollario 2.8.1: Sia $n > 1$ un numero naturale, e sia $n = p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_m^{\alpha_m}$ la sua fattorizzazione in numeri primi, dove ciascun p_i con $1 \leq i \leq m$ é un numero primo distinto, elevato ad un certo esponente α_i. L'espressione di $\varphi(n)$ puó essere anche scritta come:

$$\varphi(n) = \prod_{i=1}^{m} p_i^{\alpha_i - 1}(p_i - 1) = p_1^{\alpha_1 - 1}(p_1 - 1) p_2^{\alpha_2 - 1}(p_2 - 1) \ldots p_m^{\alpha_m - 1}(p_m - 1)$$

Dimostrazione: Questo risultato deriva direttamente dal Teorema 2.8.1. Infatti, se φ é moltiplicativa, allora:

$$\varphi(n) = \varphi(p_1^{\alpha_1})\varphi(p_2^{\alpha_2}) \ldots \varphi(p_m^{\alpha_m}) = \prod_{i=1}^{m} \varphi(p_i^{\alpha_i})$$

Applicando poi il Lemma 2.8.2 all'argomento della produttoria, si ha:

$$\prod_{i=1}^{m} \varphi(p_i^{\alpha_i}) = \prod_{i=1}^{m} p_i^{\alpha_i - 1}(p_i - 1) = p_1^{\alpha_1 - 1}(p_1 - 1)p_2^{\alpha_2 - 1}(p_2 - 1)...p_m^{\alpha_m - 1}(p_m - 1)$$

□

Il Corollario 2.8.1 permette di calcolare la funzione di Eulero in maniera molto piú semplice rispetto al calcolarla direttamente a partire dalla definizione, soprattutto per numeri molto grandi, perché richiede solamente la fattorizzazione in numeri primi e semplici moltiplicazioni.

Esempio 2.8.2: Sia $n = 246064$. La sua fattorizzazione in numeri primi é $2^4 \cdot 7 \cdot 13^3$. Si ha:

$$\varphi(246064) = \prod_{i=1}^{3} p_i^{\alpha_i - 1}(p_i - 1) = 2^{4-1}(2-1) \cdot 7^{1-1}(7-1) \cdot 13^{3-1}(13-1) = 97344$$

Teorema 2.8.2: Sia $n \in \mathbb{N}$ con $n > 0$. Il valore di $\varphi(n)$ equivale al numero di elementi invertibili di $\mathbb{Z}_n$.

2.9. Teorema di Fermat-Eulero

Teorema 2.9.1 (Piccolo Teorema di Fermat): Sia $p \in \mathbb{N}$ numero primo. Per qualsiasi $a \in \mathbb{N}$ vale:

$$a^p \equiv a \bmod p$$

Inoltre, se p non é divisore di a, vale anche:

$$a^{p-1} \equiv 1 \bmod p$$

Dimostrazione: Si consideri innanzitutto il caso in cui p non sia divisore di a. Si studino allora le classi di resto cosí definite:

$$\left\{[0]_p, [a]_p, [2a]_p, ..., [(p-1)a]_p\right\}$$

É possibile provare che tali classi sono tutte distinte fra loro. Si supponga infatti per assurdo che questo non sia vero, e che quindi esistano (almeno) due classi di resto dell'insieme sopra definito che coincidono. Siano queste $[ra]_p = [sa]_p$, con $r, s \in \mathbb{Z}$ tali per cui $0 \leq r < p$ e $0 \leq s < p$. Supponendo, senza perdita di generalitá, $r \geq s$, si ha allora:

$$[ra]_p = [sa]_p \Rightarrow [ra]_p - [sa]_p = [0]_p \Rightarrow [ra - sa]_p = [0]_p \Rightarrow [(r-s)a]_p = [0]_p$$

Ovvero, $p \mid (r-s)a$. Per il Lemma 2.4.1, deve essere vera almeno una proposizione fra $p \mid r - s$ e $p \mid a$; dato che quest'ultima non puó essere vera per ipotesi, deve aversi $p \mid r - s$. I due numeri interi r e s sono stati peró definiti come positivi ed inferiori a p, pertanto $p \mid r - s$ puó essere vera solamente nel caso in cui $r - s = 0$, ovvero $r = s$. Ma allora:

$$\left\{[0]_p, [a]_p, [2a]_p, ..., [(p-1)a]_p\right\} = \left\{[0]_p, [1]_p, [2]_p, ..., [(p-1)]_p\right\}$$

Questo perché entrambi hanno esattamente p classi di resto modulo p, e diventa allora possibile ridurre in modulo p il primo insieme ottenendo il secondo.
Poiché la classe $[0]_p$ compare in entrambi gli insiemi, puó essere eliminata matenendo valida l'uguaglianza:

$$\left\{[a]_p, [2a]_p, ..., [(p-1)a]_p\right\} = \left\{[1]_p, [2]_p, ..., [(p-1)]_p\right\}$$

Se i due insiemi sono uguali membro a membro, allora il prodotto degli elementi del primo insieme deve essere uguale al prodotto degli elementi del secondo insieme:

$$\begin{aligned}[a]_p \cdot [2a]_p \cdot \ldots \cdot [(p-1)a]_p &= [1]_p \cdot [2]_p \cdot \ldots \cdot [(p-1)]_p \Rightarrow \\ [a \cdot 2a \cdot \ldots \cdot (p-1)a]_p &= [1 \cdot 2 \cdot \ldots \cdot (p-1)]_p \Rightarrow \\ [a^{p-1}(p-1)!]_p &= [(p-1)!]_p\end{aligned}$$

Da cui si ricava, per la definizione di classe di resto, $a^{p-1}(p-1)! \equiv (p-1)! \bmod p$. Essendo p un numero primo, certamente non puó essere un divisore di $(p-1)!$, pertanto é valido il Corollario 2.6.1 e quindi é possibile semplificare come $a^{p-1} \equiv 1 \bmod p$.
Si supponga ora che p sia un numero primo qualsiasi. Se p non é un divisore di a, é possibile applicare il Corollario 2.6.1 "nell'altro verso" al risultato appena trovato. Ovvero, é possibile moltiplicare ambo i membri di $a^{p-1} \equiv 1 \bmod p$ per p, ottenendo $a^p \equiv p \bmod p$.
Se invece p é un divisore di a, questo equivale a dire $a \equiv 0 \bmod p$. Tuttavia, deve valere anche $a^p \equiv 0 \bmod p$; per proprietá transitiva, $a^p \equiv a \bmod p$. □

Teorema 2.9.2 (Teorema di Fermat-Eulero): Sia $n \in \mathbb{N} - \{0\}$ e sia a un qualsiasi intero tale che a ed n siano primi fra di loro. Allora vale:

$$a^{\varphi(n)} \equiv 1 \bmod n$$

Dimostrazione: Si consideri innanzitutto il caso in cui n sia una potenza di un numero primo, ovvero $n = p^m$ con p numero primo e m numero naturale. Si proceda per induzione su m; il caso base si ha con $m = 1$:

$$a^{\varphi(p^1)} \equiv 1 \bmod p^1 \Rightarrow a^{p^{1-1}(p-1)} \equiv 1 \bmod p \Rightarrow a^{p-1} \equiv 1 \bmod p$$

Che equivale all'enunciato del Teorema 2.9.1, e pertanto é verificato.
Si consideri ora l'ipotesi induttiva: si dimostri che sia valido $a^{\varphi(p^m)} \equiv 1 \bmod p^m$ assumendo che sia valido $a^{\varphi(p^{m-1})} \equiv 1 \bmod p^{m-1}$. Tale espressione equivale a:

$$a^{\varphi(p^{m-1})} \equiv 1 \bmod p^{m-1} \Rightarrow p^{m-1} \mid a^{\varphi(p^{m-1})} - 1 \Rightarrow a^{\varphi(p^{m-1})} - 1 = p^{m-1}b$$

Per un certo $b \in \mathbb{Z}$. Per il Lemma 2.8.2, é possibile esplicitare l'esponente di a:

$$a^{\varphi(p^{m-1})} - 1 = p^{m-1}b \Rightarrow a^{p^{m-2}(p-1)} - 1 = p^{m-1}b \Rightarrow a^{p^{m-2}(p-1)} = 1 + p^{m-1}b$$

Elevando ambo i membri alla potenza p, si ha:

$$\left(a^{p^{m-2}(p-1)}\right)^p = (1 + p^{m-1}b)^p \Rightarrow a^{p^{m-1}(p-1)} = (1 + p^{m-1}b)^p \Rightarrow a^{\varphi(p^m)} = (1 + p^{m-1}b)^p$$

Il termine $(1 + p^{m-1}b)^p$ puó essere espanso usando la formula del binomio di Newton:

$$(1 + p^{m-1}b)^p \Rightarrow 1 + (p^{m-1}b)^p + \sum_{k=1}^{p-1} \binom{p}{k} (p^{m-1}b)^{p-k}$$

Ogni addendo della sommatoria, cioè ogni termine $\binom{p}{k}(p^{m-1}b)^{p-k}$, é un multiplo di p^m perché $\binom{p}{k}$ é multiplo di p e $(p^{m-1}b)^{p-k}$ é multiplo di p^{m-1}, per $k = 1, \ldots, p-1$.
Inoltre, $(p^{m-1}b)^p$ è un multiplo di p^m, dunque si ha:

$$(1 + p^{m-1}b)^p \equiv 1 \bmod p^m$$

Da cui, per proprietá transitiva:

$$a^{\varphi(p^m)} \equiv 1 \bmod p^m$$

Nel caso in cui n sia un numero qualsiasi, questo puó essere certamente fattorizzato come $n = p_1^{m_1} p_2^{m_2} ... p_r^{m_r}$, dove ciascun p_i con $1 \leq i \leq r$ é un numero primo distinto e ciascun m_i é un numero naturale. Per ciascuno di questi fattori elevati al rispettivo esponente, dovrá valere:

$$a^{\varphi\left(p_i^{m_i}\right)} \equiv 1 \bmod p_i^{m_i}$$

Per il Teorema 2.8.1, si ha che ciascun $\varphi\left(p_i^{m_i}\right)$ é divisore di $\varphi(n)$, ovvero che per un certo $t \in \mathbb{Z}$ vale $\varphi(n) = \varphi\left(p_i^{m_i}\right)t$. Allora:

$$a^{\varphi(n)} = a^{\varphi\left(p_i^{m_i}\right)t} = \left(a^{\varphi\left(p_i^{m_i}\right)}\right)^t \equiv 1^t = 1 \bmod p_i^{m_i}$$

In altre parole, ogni $p_i^{m_i}$ é divisore di $a^{\varphi(n)} - 1$. Dato che ogni $p_i^{m_i}$ é potenza di un numero primo, é evidente come, presi due $p_i^{m_i}$ e $p_j^{m_j}$ qualsiasi con $i \neq j$, questi saranno coprimi. Ma allora:

$$p_1^{m_1} p_2^{m_2} ... p_r^{m_r} \mid a^{\varphi(n)} - 1 \Rightarrow a^{\varphi(n)} \equiv 1 \bmod p_1^{m_1} p_2^{m_2} ... p_r^{m_r}$$

Avendo peró definito n come $n = p_1^{m_1} p_2^{m_2} ... p_r^{m_r}$:

$$a^{\varphi(n)} \equiv 1 \bmod n$$

□

Calcolare $a^m \bmod n$ "a mano" puó richiedere molto tempo, specialmente se i numeri in questione sono molto grandi. É peró possibile velocizzare il procedimento impiegando il **metodo dei quadrati ripetuti**, di seguito discusso.
Si scriva l'esponente m in base due, ottenendo $m = \sum_{i=0}^{k-1} d_i 2^i$, dove ciascun d_i é la i-esima cifra della rappresentazione. Si costruisca poi una tabella come quella seguente:

$$\left[\begin{array}{c|c} & c_0 = 1 \\ d_{k-1} & c_1 \equiv c_0^2 \cdot a^{d_{k-1}} \bmod n \\ d_{k-2} & c_2 \equiv c_1^2 \cdot a^{d_{k-2}} \bmod n \\ \vdots & \vdots \\ d_1 & c_{k-1} \equiv c_{k-2}^2 \cdot a^{d_1} \bmod n \\ d_0 & c_k \equiv c_{k-1}^2 \cdot a^{d_0} \bmod n \end{array}\right]$$

Risulta $a^m \equiv c_k \bmod n$.

Esempio 2.9.1: Si voglia calcolare $3^{90} \bmod 91$. Si ha $90_{10} = 1011010_2$. Si ha:

$$\left[\begin{array}{c|l} & c_0 = 1 \\ 1 & c_1 \equiv 1^2 \cdot 3^1 = 3 \bmod 91 \\ 0 & c_2 \equiv 3^2 \cdot 3^0 = 9 \bmod 91 \\ 1 & c_3 \equiv 9^2 \cdot 3^1 = 273 \equiv -30 \bmod 91 \\ 1 & c_4 \equiv (-30)^2 \cdot 3^1 = 2700 \equiv -30 \bmod 91 \\ 0 & c_5 \equiv (-30)^2 \cdot 3^0 = 900 \equiv -10 \bmod 91 \\ 1 & c_6 \equiv (-10)^2 \cdot 3^1 = 300 \equiv 27 \bmod 91 \\ 0 & c_7 \equiv (27)^2 \cdot 3^0 = 729 \equiv 1 \bmod 91 \end{array}\right]$$

Risulta $3^{90} \equiv 1 \bmod 91$

2.10. Test di primalitá

Si voglia generare, fissato un certo ordine di grandezza, un numero primo qualsiasi di tale dimensione. L'approccio piú semplice consiste nello scegliere un numero n dispari di tale ordine di grandezza fissato e valutare se

tale numero é primo; se non lo é, si considera $n+2$ e si valuta se é primo, se non lo é si valuta $n+4$, ecc... Per il Teorema dei Numeri Primi, é garantito che un numero venga trovato entro $O(\log(n))$ passi.
Viene detto **test di primalitá** un procedimento, in genere espresso sotto forma di algoritmo, che permette di determinare se un numero intero n qualsiasi sia o non sia un numero primo.
Il test piú semplice, detto *test naive*, prevede di calcolare, per ciascun $1 < k < n$, la divisione fra n e k: se esiste almeno un k tale per cui $k \mid n$, allora n non é primo, altrimenti lo é. Questo approccio puó essere migliorato osservando come nell'intervallo $1 < k < n$ possono ripetersi piú volte dei multipli di numeri primi: se tali numeri fossero divisori di n, per il Lemma 2.4.1 anche i loro fattori lo sarebbero, ma questi sarebbero giá stati testati (essendo certamente minori di n). Pertanto, non é necessario considerare tutti i numeri nell'intervallo $1 < k < n$, ma soltanto quelli primi.

Lemma 2.10.1: Sia $n \in \mathbb{N}$ con $n \geq 1$. Se n é un numero composto, allora almeno uno dei numeri primi che costituiscono la sua fattorizzazione é minore di $\lfloor \sqrt{n} \rfloor$.

Dimostrazione: Si supponga che, in una certa iterazione del test naive, venga trovato un numero p che é divisore di n. Allora é possibile scrivere $n = pq$ per un certo $q \in \mathbb{Z}$. Per come il test é strutturato, q deve necessariamente essere maggiore o uguale a p, perché altrimenti il test lo avrebbe giá individuato (o avrebbe individuato un numero primo della sua fattorizzazione). Se q fosse maggiore di p, allora non sarebbe rilevante ai fini del test, perché p verrebbe scoperto prima di q ed il test terminerebbe comunque. L'unico caso rilevante si ha con $p = q$, ovvero $n = p \cdot p = p^2$, da cui si ha $\sqrt{n} = p$. Dato che p é stato scelto casualmente, si ha che deve esistere almeno un p con queste caratteristiche nell'intervallo $(1, n)$. □

Il test naive é un test *deterministico*, ovvero garantisce di restituire sempre la risposta corretta. In altre parole, se il test determina che un certo numero n é un numero primo, allora tale numero é effettivamente un numero primo, mentre se determina che é un numero composto allora é effettivamente un numero composto. Il problema di tale test é che richiede troppe computazioni, specialmente per numeri grandi.
Approcci diversi sono forniti dai test *probabilistici*, ovvero che non garantiscono di fornire la risposta corretta ma lo fanno a meno di una certa percentuale. Il vantaggio di tali test é che sono molto piú veloci dei test deterministici, e quindi utilizzabili nella pratica (specialmente quando un certo margine di errore é tollerato). Inoltre, piú test probabilistici possono essere applicati ad uno stesso numero: piú test confermano lo stesso risultato e maggiore é la certezza del responso.
Siano $n > 1$ un numero intero dispari e b un intero qualsiasi, primo con n. Se vale $b^{n-1} \equiv 1 \bmod n$ si dice che n è **pseudoprimo di Fermat** rispetto alla base b. La locuzione "pseudoprimo di Fermat" viene dal fatto che la definizione di pseudoprimo dipende dal contesto; se non diversamente specificato, ci si riferirá agli pseudoprimi di Fermat semplicemente come pseudoprimi.

Esempio 2.10.1: 15 é uno pseudoprimo per la base 4. Infatti, 4 e 15 sono primi fra di loro ed é vero che $4^{14} \equiv 1 \bmod 15$.

Lemma 2.10.2: Un numero primo p é pseudoprimo rispetto a qualsiasi base.

Teorema 2.10.1: Per ogni intero $b > 1$, esistono infiniti pseudoprimi rispetto alla base b.

Alla luce dei risultati trovati, é possibile enunciare un semplice algoritmo, detto **test di primalitá di Fermat**, che determina se un numero intero n é o non é un numero primo. Se n é pari ed é diverso da 2, allora é certamente un numero composto. Pertanto, senza perdita di generalitá, si assuma che n sia dispari:

1. Si fissi un parametro k, che determina il numero di volte che l'algoritmo verrá eseguito;
2. Si scelga un qualsiasi numero b tale per cui $0 < b < n$;
3. Si calcoli $\text{MCD}(b, n)$ con l'algoritmo di Euclide;
4. Se $\text{MCD}(b, n) > 1$, allora n é certamente un numero composto, perché ha almeno $\text{MCD}(b, n)$ come divisore, e l'algoritmo termina. Altrimenti, si calcoli $b^{n-1} \bmod n$;
5. Se $b^{n-1} \not\equiv 1 \bmod n$, allora n é certamente un numero composto, perché altrimenti violerebbe il Teorema 2.9.1, e l'algoritmo termina. Se invece $b^{n-1} \equiv 1 \bmod n$, l'iterazione corrente per il test é "inconclusiva";
6. Se l'algoritmo é giá stato eseguito k volte, allora n é *probabilmente* un numero primo, e l'algoritmo termina, altrimenti riprende dal punto 2.

L'algoritmo garantisce di determinare che un numero sia un numero composto se é un numero composto, ma non da garanzie di determinare che un numero sia primo se é un numero primo.
Si supponga di aver applicato l'algoritmo k volte, usando quindi k basi $b_1, b_2, ..., b_k$, e di aver trovato che n é probabilmente primo. L'efficienza del test dipende dalla probabilitá che n sia effettivamente primo.

Teorema 2.10.2: Sia $n > 1$ un intero composto dispari. Se n non è pseudoprimo rispetto ad almeno una base b, allora n non è pseudoprimo per almeno la metà delle basi possibili viste modulo n, cioè le $\varphi(n)$ basi b con $0 < b < n$ e $\text{MCD}(b, n) = 1$.

Il Teorema 2.10.2 permette di dare una stima della probabilitá che un numero n che il test di Fermat stabilisce essere primo sia effettivamente primo.
Se n è un numero composto e vale $b_1^{n-1} \equiv 1 \bmod n$, si ha che n é pseudoprimo rispetto a b_1. Per il Teorema 2.10.2, la probabilitá che n "superi" il test pur non essendo un numero primo é $\frac{1}{2}$. Dato che ogni iterazione dell'algoritmo é indipendente dalle altre, la probabilitá che n "superi" tutte e k le iterazioni del test pur non essendo un numero primo é $\frac{1}{2^k}$.

3. Gruppi

3.1. Proprietá dei gruppi

Per comoditá, verranno fatte delle semplificazioni di notazione. Se non riportato diversamente:

- L'operazione che figura nei gruppi, a prescindere da quale sia, verrá denotata con $\diamond$;
- Per indicare $x \diamond y$ (con $x, y \in G$) verrá usata la notazione abbreviata xy;
- L'elemento neutro per l'operazione $\diamond$ verrá indicato con 1;
- L'applicazione reiterata k volte dell'operazione $\diamond$ su un $x \in G$, ovvero $\underbrace{x \diamond x \diamond ... \diamond x}_{k}$, verrá indicata con x^k;
- Il reciproco di un $x \in G$ rispetto a $\diamond$ viene indicato con $\overline{x}$;
- Se ci si sta riferendo ad un gruppo ed é noto dal contesto quale sia l'operazione che figura nel gruppo, ci si riferirá al gruppo solo con il suo insieme sostegno. In altre parole, se $(G, \diamond)$ é un gruppo ed é noto dal contesto che l'operazione a cui ci si riferisce é $\diamond$, si indicherá con il solo G la coppia $(G, \diamond)$.

Lemma 3.1.1: Sia G un gruppo. Per qualsiasi $x, y, z \in G$, vale:

- Unicitá del reciproco: $\exists! x^{-1} : x \diamond x^{-1} = 1$;
- Unicitá dell'elemento neutro: $\exists! 1 : x \diamond 1 = x$;
- Legge di cancellazione (a destra): $x \diamond y = x \diamond z \Rightarrow y = z$;
- Legge di cancellazione (a sinistra): $y \diamond x = z \diamond x \Rightarrow y = z$.

Dato un gruppo G, la struttura algebrica $(H, \diamond)$ si dice **sottogruppo** di G se H é un sottoinsieme (anche improprio) di G e se la coppia $(H, \diamond)$ forma a sua volta un gruppo. In altre parole, $H = (H, \diamond)$ é un sottogruppo di $G = (G, \diamond)$ se:

- L'elemento neutro di G appartiene ad H;
- L'insieme H é *chiuso* rispetto all'operazione $\diamond$, ovvero $\forall h, k \in H$ si ha $h \diamond k \in H$;
- $\forall h \in H$, il reciproco $\overline{h}$ di h é a sua volta membro di H.

Per indicare che H é un sottogruppo di G si usa la notazione $H \leq G$. Se H é un sottogruppo di G ed é distinto da G si dice che H é un **sottogruppo proprio** di G, e si indica con $H < G$.
Si noti come le notazioni $<$ e $\leq$ non hanno nulla a che vedere con le relazioni d'ordine "minore" e "minore o uguale" rispetto ai numeri, cosí come non si riferiscono alla cardinalitá dei sostegni dei gruppi. Infatti, é accettato che due gruppi possano essere l'uno il sottogruppo dell'altro pur avendo la stessa cardinalitá.

Lemma 3.1.2: Sia $G = (G, \diamond)$ un gruppo. Un sottoinsieme H di G é un sottogruppo di G se e soltanto se, per ogni coppia di elementi (non necessariamente distinti) $h, k \in H$, vale $h \diamond \overline{k} \in H$.

Dimostrazione: Se é noto che H sia un sottogruppo di G, allora H rispetta certamente la proprietá richiesta. Infatti, se $(H, \diamond)$ é un gruppo, allora é chiuso rispetto a $\diamond$, e quindi $\forall h, k \in H$ vale $h \diamond k \in H$. Inoltre, $\forall h \in H, \overline{h} \in H$, pertanto $\overline{k} \in H$, e si ha quindi $h \diamond \overline{k} \in H$ per ogni $h, k \in H$.
Viceversa, si supponga che H sia un sottoinsieme di G tale per cui $\forall h, k \in H$ vale $h \diamond \overline{k} \in H$:

- Se $h = k$, allora, per l'unicitá del reciproco, $\overline{h} = \overline{k}$, e quindi $h \diamond \overline{h} = h \diamond \overline{k} = 1$, quindi l'elemento neutro di $(G, \diamond)$ appartiene ad H;
- Se $h = 1$ (ed é lecito, avendo appena mostrato che appartiene ad H), allora per un qualsiasi k vale $1 \diamond \overline{k} \in H$, ma $1 \diamond \overline{k} = \overline{k}$ per definizione di elemento neutro. Si ha quindi che $\forall h \in H$, vale $\overline{h} \in H$;
- Siano $h, k \in H$. Avendo appena provato che $\overline{k}$ appartiene ad H per un qualsiasi $k \in H$, vale $h \diamond \overline{\overline{k}} \in H$, ma $\overline{\overline{k}} = k$, pertanto $h \diamond k \in H$.

Si ha quindi che H rispetta la definizione di sottogruppo, pertanto $H \leq G$. □

Il Lemma 3.1.2 é un possibile criterio che permette di determinare se, dati due gruppi G ed H, H sia un sottogruppo di G.

Esempio 3.1.1: La struttura algebrica $(\mathbb{Z}, +)$ é un gruppo. La struttura algebrica $(n\mathbb{Z}, +)$, dove $n\mathbb{Z} = \{nk : k \in \mathbb{Z}\}$ é l'insieme che contiene tutti i multipli (interi) di n, é un sottogruppo di $(\mathbb{Z}, +)$. Siano infatti a e b due elementi di $(n\mathbb{Z})$. Si ha:

$$a + \overline{b} = nk_1 + \overline{nk_2} = nk_1 - nk_2 = n(k_1 - k_2)$$

Dato che $(k_1 - k_2) \in \mathbb{Z}$, si ha $n(k_1 - k_2) \in n\mathbb{Z}$. Pertanto, per il Lemma 3.1.2 si ha che $(n\mathbb{Z}, +)$ é sottogruppo di $(\mathbb{Z}, +)$. Si noti inoltre come i due insiemi sostegno, $\mathbb{Z}$ e $n\mathbb{Z}$, abbiano la stessa cardinalitá.

Lemma 3.1.3: Per un qualsiasi gruppo $G = (G, \diamond)$, le strutture algebriche $(G, \diamond)$ e $(\{1\}, \diamond)$ sono sottogruppi di G.

Dimostrazione:
- L'insieme G della struttura algebrica $(G, \diamond)$ é lo stesso insieme che figura nell'insieme G del gruppo $G = (G, \diamond)$. Pertanto, il Lemma 3.1.2 é certamente verificato;
- L'unico elemento che figura nell'insieme $\{1\}$ della struttura algebrica $(\{1\}, \diamond)$ é precisamente 1. A prescindere di come $\diamond$ sia definita, si ha $\overline{1} = 1$, pertanto $1 \diamond \overline{1} = 1 \diamond 1 = 1$. Dato che $1 \in \{1\}$, il Lemma 3.1.2 é verificato.

□

Per un qualsiasi gruppo G, il sottogruppo G viene detto **sottogruppo improprio**, mentre il sottogruppo $\{1\}$ viene detto **sottogruppo banale**.

3.2. Permutazioni

Sia X un insieme. Una funzione biettiva $\sigma : X \mapsto X$ si dice **permutazione** su X. L'insieme di tutte le permutazioni che é possibile costruire per X viene indicato con S_X.

Lemma 3.2.1: Sia X un insieme e sia S_X l'insieme di tutte le permutazioni costruibili per X. Se X é un insieme finito di cardinalitá n, allora $|S_X| = n!$.

Dimostrazione: Se σ é una permutazione su X e $X = \{x_1, ..., x_n\}$, allora esistono n possibili scelte per l'immagine $\sigma(x_1)$. Scelto poi un secondo elemento $x_2 \neq x_1$, questo avrá $n - 1$ scelte per $\sigma(x_2)$, perché σ é per definizione iniettiva (essendo biettiva, quindi iniettiva e suriettiva) ed una delle scelte é giá occupata da x_1. Ripetendo questo ragionamento per tutti gli elementi di X, si ha che il numero di permutazioni su X é esattamente $n \cdot (n-1) \cdot ... \cdot 1 = n!$. □

Teorema 3.2.1: Sia X un insieme e sia S_X l'insieme di tutte le permutazioni costruibili per X. La struttura algebrica $(S_X, \circ)$, dove $\circ$ é l'operazione di composizione di funzioni, costituisce un gruppo.

Dimostrazione: La struttura algebrica $(S_X, \circ)$ forma un semigruppo perché, per il Teorema 1.3.1, l'operazione di composizione gode della proprietá associativa. É inoltre un monoide, perché l'operazione di composizione ha nella funzione identitá l'elemento neutro, come da Corollario 1.3.1. É infine un gruppo perché, essendo biettiva per definizione, per ogni permutazione ne esiste una inversa, e la funzione inversa é l'inverso rispetto alla composizione, come da Corollario 1.3.2. □

Per un insieme X, il gruppo $(S_X, \circ)$ viene chiamato **gruppo simmetrico** o **gruppo delle permutazioni**.

Essendo $\circ$ l'operazione piú "interessante" da applicare alle permutazioni, si usa chiamare *prodotto* di due permutazioni la loro composizione. Pertanto, se σ e τ sono due permutazioni in S_X, la scrittura $\sigma \circ \tau$ puó anche venire riportata come $\sigma\tau$.
In genere, quando si parla di permutazioni su un insieme X, si ha interesse a considerare X come i primi n numeri interi, ovvero come $X = \{1, 2, ..., n\}$. Per tal motivo, viene usata la notazione S_n per indicare l'insieme di tutte le permutazioni su $X = \{1, 2, ..., n\}$, sottointentendo che l'insieme a cui S_n si riferisce sia quest'ultimo.
Una certa permutazione $\sigma \in S_n$ viene spesso indicata anche come:

$$\sigma = \begin{pmatrix} x_1 & x_2 & \dots & x_i & \dots & x_n \\ y_1 & y_2 & \dots & y_i & \dots & y_n \end{pmatrix}$$

Dove, per ogni i, si ha $y_i = \sigma(x_i)$. L'ordinamento della prima riga puó essere arbitrario, ma per convenzione viene in genere ordinata in ordine crescente.

Esempio 3.2.1:
- Con $n = 3$, si hanno $3! = 6$ permutazioni possibili, che possono pertanto essere facilmente enumerate:

$$S_3 = \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \right\}$$

- Con $n = 12$, si hanno $12! = 479001600$ permutazioni possibili. Una di queste é:

$$\sigma \in S_{12} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ 8 & 1 & 2 & 7 & 12 & 5 & 9 & 3 & 11 & 4 & 6 & 10 \end{pmatrix}$$

Si dice che la permutazione $\sigma \in S_n$ *muove* un elemento a se $\sigma(a) \neq a$, ovvero se "sposta" l'elemento a in una posizione diversa da quella in cui si trova. In caso contrario, ovvero se $\sigma(a) = a$, si dice che σ *fissa* a.
L'insieme costituito dagli elementi mossi da σ prende il nome di **supporto** di σ. Due permutazioni $\sigma, \tau \in S_n$ si dicono **disgiunte** se i loro supporti sono insiemi disgiunti.

Esempio 3.2.2: Si considerino le tre permutazioni $\sigma, \tau, \upsilon \in S_6$:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 6 & 2 & 3 & 5 & 4 \end{pmatrix} \qquad \tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 2 & 3 & 4 & 1 & 6 \end{pmatrix} \qquad \upsilon = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 3 & 5 & 6 & 4 \end{pmatrix}$$

Il supporto di σ é $\{2, 3, 4, 6\}$, quello di τ é $\{1, 5\}$ mentre quello di υ é $\{4, 5, 6\}$. Si ha quindi che σ e τ sono disgiunte.

Teorema 3.2.2: Se σ e τ sono due permutazioni disgiunte, si ha $\sigma\tau = \tau\sigma$.

Una permutazione nella forma:

$$\begin{pmatrix} x_1 & x_2 & \dots & x_{r-1} & x_r & x_{r+1} & \dots & x_n \\ x_2 & x_3 & \dots & x_r & x_1 & x_{r+1} & \dots & x_n \end{pmatrix}$$

Viene detta **permutazione ciclica** di lunghezza r, o semplicemente **ciclo** di lunghezza r, con $r \geq 2$. Intuitivamente, un ciclo "sposta" l'elemento x_b "sotto" a x_a, l'elemento x_b "sotto" a x_c, ..., e x_r "sotto" a x_a.
Dato che un ciclo o fissa o muove ciascun suo elemento, per denotare un ciclo é sufficiente denotare quali elementi vengono mossi ed in quale posizione, perché tutti gli elementi non menzionati sono implicitamente fissati.
Un ciclo σ di lunghezza r viene denotato con $\sigma = (x_1, x_2, ..., x_r)$; tale scrittura sta ad indicare che in corrispondenza di ciascun elemento x_i viene messo l'elemento x_{i+1}, ad eccezione dell'r-esimo elemento che viene messo in corrispondenza con x_1.
Si noti come la scrittura $(x_1, x_2, x_3, ..., x_r)$ sia equivalente alla scrittura $(x_r, x_1, x_2, ..., x_{r-1})$ e alla scrittura $(x_{r-1}, x_r, x_1, ..., x_{r-2})$, ecc... perché sono tutti cicli che inducono il medesimo "spostamento", semplicemente

si prende come "riferimento iniziale" un suo elemento diverso. Nello specifico, ogni ciclo puó essere scritto in tanti modi diversi quant'é la sua lunghezza.

Esempio 3.2.3: Quello presentato di seguito é un ciclo di lunghezza 5, appartenente all'insieme delle permutazioni S_{12}:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ 9 & 2 & 3 & 6 & 5 & 11 & 7 & 8 & 4 & 10 & 1 & 12 \end{pmatrix}$$

Tale ciclo mette 1 in corrispondenza con 9, 9 corrispondenza con 4, 4 in corrispondenza con 6, 6 in corrispondenza con 11 e 11 in corrispondenza con 1. Pertanto, viene denotato come $(1, 9, 4, 6, 11)$. Si noti come tale scrittura possa essere formulata in 5 modi, tutti equivalenti:

$$(1,9,4,6,11) \qquad (11,1,9,4,6) \qquad (6,11,1,9,4) \qquad (4,6,11,1,9) \qquad (9,4,6,11,1)$$

Teorema 3.2.3: Ogni permutazione di S_n, diversa dalla identità, è un ciclo oppure è il prodotto di cicli disgiunti, univocamente determinati a meno dell'ordine.

Esempio 3.2.4: La permutazione $\sigma \in S_{13}$ a sinistra puó essere scomposta nel prodotto dei tre cicli $\upsilon_1, \upsilon_2, \upsilon_3$ a destra:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 \\ 9 & 12 & 13 & 6 & 7 & 11 & 2 & 3 & 4 & 10 & 1 & 5 & 8 \end{pmatrix} \qquad \begin{aligned} \upsilon_1 &= (1,9,4,6,11), \\ \upsilon_2 &= (2,12,5,7), \\ \upsilon_3 &= (3,13,8) \end{aligned}$$

Per convincersene, é sufficiente comporre (in ordine arbitrario) i tre cicli. Si consideri, per esempio, $\upsilon_1 \circ \upsilon_2 \circ \upsilon_3$:

$$\upsilon_1 \circ \upsilon_2 \circ \upsilon_3 = \upsilon_1\upsilon_2\upsilon_3 = \upsilon_1(\upsilon_2(\upsilon_3)) = \upsilon_1\left(\upsilon_2\left(\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 \\ 1 & 2 & 13 & 4 & 5 & 6 & 7 & 3 & 9 & 10 & 11 & 12 & 8 \end{pmatrix}\right)\right)$$

$$= \upsilon_1\left(\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 \\ 1 & 12 & 13 & 4 & 7 & 6 & 2 & 3 & 9 & 10 & 11 & 5 & 8 \end{pmatrix}\right) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 \\ 9 & 12 & 13 & 6 & 7 & 11 & 2 & 3 & 4 & 10 & 1 & 5 & 8 \end{pmatrix}$$

3.3. Polinomi su un campo

Sia dato un certo campo K. Prende il nome di **polinomio** a coefficienti in K e incognita in x qualunque espressione nella forma:

$$p(x) = \sum_{i=0}^{n} a_i x^i = a_n x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0 \quad \text{con} \quad n \in \mathbb{N}, a_i \in K \ \forall i \in \{0, ..., n\}$$

Dove l'intero non negativo n si dice **grado** di $p(x)$ e lo si indica con $\partial p(x)$. Al polinomio nullo $p(x) = 0$ si attribusce, per convenzione, grado -1. Il coefficiente a_n si dice **coefficiente direttore** di $p(x)$. Se $a_n = 1$, si dice che $p(x)$ è **monico**.
Sui polinomi é possibile definire delle operazioni di somma e di prodotto. Siano $p(x)$ e $q(x)$ due polinomi a coefficienti in un campo K, di grado rispettivamente n e m. Assumendo, senza perdita di generalitá, che $n \geq m$, la somma fra $p(x)$ e $q(x)$ viene definita come:

$$p(x) + q(x) = \sum_{i=\min(n,m)+1}^{n} a_i x^i + \sum_{i=0}^{\min(n,m)} (a_i + b_i)x^i$$

Mentre il prodotto come:

$$p(x)q(x) = \sum_{0}^{n+m} \left(\sum_{i+j=k} a_i b_i \right) x^i$$

Lemma 3.3.1: La struttura algebrica $K[x]$, che ha per sostegno l'insieme di tutti i polinomi a coefficienti in K e incognita in x e per operazioni la somma fra polinomi ed il prodotto fra polinomi, é un anello commutativo. L'elemento neutro per la somma é il polinomio $p(x) = 0$, mentre l'elemento neutro per il prodotto é $p(x) = 1$.

Per l'anello $K[x]$ é possibile sviluppare una teoria parallela a quella dell'anello $\mathbb{Z}$

Teorema 3.3.1 (Algoritmo della divisione per polinomi): Siano $a(x), b(x) \in K[x]$, con $b(x)$ non nullo. Esiste una ed una sola coppia di polinomi $q(x), r(x) \in K[x]$ tali che:

1. $a(x) = b(x)q(x) + r(x)$
2. $\partial r(x) < \partial b(x)$

Dati $a(x), b(x) \in K[x]$, i due polinomi $q(x), r(x) \in K[x]$ che figurano nel Teorema 3.3.1 sono chiamati rispettivamente **quoziente** e **resto** della divisione fra $a(x)$ e $b(x)$. Se $r(x) = 0$, si dice che $b(x)$ *divide* $a(x)$, e si indica con $b(x) \mid a(x)$; se invece $b(x)$ non divide $a(x)$, si indica con $b(x) \nmid a(x)$.
Il Teorema 3.3.1 fornisce implicitamente un algoritmo che permette di calcolare la divisione fra due polinomi.

Esempio 3.3.1: Siano $a(x) = x^3 - 2x^2 + x - 1$ e $b(x) = 2x^2 - 5$ polinomi sull'anello $\mathbb{Q}[x]$. Si ha:

$$\begin{array}{rrrr|rr} x^3 & -2x^2 & +x & -1 & 2x^2 & -5 \\ -x^3 & & +\frac{5}{2}x & & & \\ \hline & -2x^2 & +\frac{7}{2}x & -1 & & \\ & 2x^2 & & -5 & & \\ \hline & & +\frac{7}{2}x & -6 & \frac{1}{2}x & -1 \end{array} \qquad p(x) = \left(\frac{1}{2}x - 1\right) \qquad q(x) = \left(\frac{7}{2}x - 6\right)$$

Siano $a(x) = [2]_7 x^4 + [-1]_7 x^2 + [1]_7$ e $b(x) = [3]_7 x^3 + [-2]_7$ polinomi sull'anello $\mathbb{Z}_7[x]$. Si ha:

$$\begin{array}{rrrr|rr} [2]_7 x^4 & +[-1]_7 x^2 & & +[1]_7 & [3]_7 x^3 & +[-2]_7 \\ [-2]_7 x^4 & & +[6]_7 x & & & \\ \hline & [-1]_7 x^2 & +[6]_7 x & +[1]_7 & [3]_7 x & \end{array} \qquad p(x) = [3]_7 x \qquad q(x) = [-1]_7 x^2 + [6]_7 x + [1]_7$$

Siano $a(x)$ e $b(x)$ due polinomi non nulli in $K[x]$. Si dice **massimo comun divisore** tra $a(x)$ e $b(x)$ ogni polinomio $d(x)$ in $K[x]$ tale che:

1. $d(x) \mid a(x)$ e $d(x) \mid b(x)$;
2. Se $c(x) \in K[x]$ tale per cui $c(x) \mid a(x)$ e $c(x) \mid b(x)$, allora $c(x) \mid d(x)$.

Teorema 3.3.2 (Esistenza di un massimo comun divisore per i polinomi): Per qualsiasi $a(x), b(x) \in K[x]$ esiste sempre un massimo comun divisore $d(x)$ fra $a(x)$ e $b(x)$. Esistono inoltre due polinomi $f(x), g(x) \in K[x]$ tali per cui:

$$a(x)f(x) + b(x)g(x) = d(x)$$

Che non é altro che l'identitá di Bézout rispetto ai polinomi.

Il massimo comun divisore tra due polinomi è determinato a meno di una costante moltiplicativa non nulla.

Lemma 3.3.2: Siano $a(x)$ e $b(x)$ due polinomi su $K[x]$, e sia $d(x)$ un massimo comun divisore fra $a(x)$ e $b(x)$. Allora $\tilde{d}(x)$ è un massimo comun divisore tra $a(x)$ e $b(x)$ se e soltanto se $\tilde{d}(x) = kd(x)$ con $k \in K - \{0\}$.

Corollario 3.3.1: Dati due polinomi $a(x)$ e $b(x)$ su $K[x]$, esiste uno ed un solo polinomio monico $d(x)$ che sia massimo comun divisore tra $a(x)$ e $b(x)$.

Dimostrazione: Se per il Lemma 3.3.2 i massimi comuni divisori fra due polinomi sono determinati a meno di una costante, allora esiste un solo polinomio che abbia 1 come coefficiente direttore, ovvero un solo polinomio monico. □

Per comoditá, con $\mathrm{MCD}(a(x), b(x))$ si indica il massimo comun divisore fra i polinomi $a(x)$ e $b(x)$ su $K[x]$ che ha 1 come coefficiente direttore. In particolare, se il grado del massimo comun divisore è zero, allora tale massimo comun divisore è 1. In questo caso, i polinomi $a(x)$ e $b(x)$ si dicono **coprimi** o **primi fra di loro**.

Esempio 3.3.2:
Siano $a(x) = x^3 + 1$ e $b(x) = x^2 + 1$ polinomi sull'anello $\mathbb{Q}[x]$. Si ha:

$$\begin{array}{rrr|rr} x^3 & & +1 & x^2 & +1 \\ \hline -x^3 & -x & & & \\ \hline & -x & +1 & x & \end{array} \qquad \begin{array}{rrr|rr} x^2 & & +1 & -x & +1 \\ \hline -x^2 & +x & & & \\ \hline & x & +1 & & \\ & -x & +1 & & \\ \hline & & 2 & -x & -1 \end{array}$$

$$a(x) = b(x)(x) + (-x+1)$$
$$b(x) = (-x+1)(-x-1) + 2$$

Un massimo comun divisore fra $a(x)$ e $b(x)$ é 2, pertanto $\text{MCD}(a(x), b(x)) = 1$. É poi possibile costruire l'identitá di Bézout come:

$$\begin{aligned} a(x) = b(x)(x) + (-x+1) &\Rightarrow (-x+1) = a(x) - b(x)(x) \\ b(x) = (-x+1)(-x-1) + 2 &\Rightarrow b(x) - [a(x) - b(x)(x)](-x-1) = 2 \Rightarrow \\ & b(x) + a(x)x + a(x) - b(x)(x^2) - b(x)(x) = 2 \Rightarrow \\ & a(x)(x+1) + b(x)(-x^2 - x + 1) = 2 \Rightarrow \\ & a(x)\left(\frac{x}{2} + \frac{1}{2}\right) + b(x)\left(-\frac{x^2}{2} - \frac{x}{2} + \frac{1}{2}\right) = 1 \end{aligned}$$

Siano $a(x) = [1]_5 x^3 + [1]_5 x^2 + [1]_5 x + [1]_5$ e $b(x) = [3]_5 x^2 + [2]_5 x + [2]_5$ polinomi sull'anello $\mathbb{Z}_5[x]$. Si ha:

$$\begin{array}{rrrr|rrr} [1]_5 x^3 & +[1]_5 x^2 & +[1]_5 x & +[1]_5 & [3]_5 x^2 & +[2]_5 x & +[2]_5 \\ [-6]_5 x^3 & +[-4]_5 x^2 & +[-4]_5 x & & & & \\ \hline & [2]_5 x^2 & +[2]_5 x & +[1]_5 & & & \\ & [-12]_5 x^2 & +[-8]_5 x & +[-8]_5 & & & \\ \hline & & [4]_5 x & +[3]_5 & [2]_5 x & +[4]_5 & \end{array} \qquad \begin{array}{rrr|rr} [3]_5 x^2 & +[2]_5 x & +[2]_5 & [4]_5 x & +[3]_5 \\ [-8]_5 x^2 & +[-6]_5 x & & & \\ \hline & [1]_5 x & +[2]_5 & & \\ & [-16]_5 x & +[-12]_5 & & \\ \hline & & 0 & [2]_5 x & +[4]_5 \end{array}$$

$$a(x) = b(x)\big([2]_5 x + [4]_5\big) + \big([4]_5 x + [3]_5\big) \Rightarrow a(x)[1]_5 - b(x)\big([2]_5 x + [4]_5\big) = [4]_5 x + [3]_5$$

Si ha quindi che un massimo comun divisore fra $a(x)$ e $b(x)$ é $[4]_5 x + [3]_5$. Il massimo comun divisore di $a(x)$ e $b(x)$ che ha $[1]_5$ come coefficiente direttore, ovvero $\text{MCD}(a(x), b(x))$, si ottiene dividendo $[4]_5 x + [3]_5$ per $[4]_5$, ovvero $[1]_5 x + [2]_5$.

$$a(x)[4]_5 + b(x)\big([2]_5 x + [4]_5\big) = [1]_5 x + [2]_5$$

Sia $p(x)$ un polinomio in $K[x]$, con $\partial p(x) > 0$. Il polinomio $p(x)$ si dice **primo** se, per qualsiasi $a(x), b(x) \in K[x]$, $p(x) \mid a(x)b(x)$ implica $p(x) \mid a(x)$ oppure $p(x) \mid b(x)$.
Il polinomio $p(x) \in K[x]$ con $\partial p(x) > 0$ viene detto **irriducibile** se i suoi divisori sono solo e soltanto i polinomi di grado 0 ed i polinomi nella forma $hp(x)$, con $h \in (K - \{0\})$.

Teorema 3.3.3: Il polinomio $p(x) \in K[x]$, con $\partial p(x) > 0$ é primo se e solo se é irriducibile (ovvero, le due definizioni sono equivalenti).

Si dice che un polinomio $p(x) \in K[x]$ viene **fattorizzato in polinomi primi** quando tale polinomio viene scritto come prodotto di soli polinomi primi (non necessariamente distinti) appartenenti a $K[x]$. In genere, una fattorizzazione viene espressa raccogliendo a fattor comune i polinomi primi per mettere in evidenza la loro molteplicitá. Naturalmente, la fattorizzazione in polinomi primi di un polinomio primo é sé stesso, a meno di una costante moltiplicativa non nulla.

Esempio 3.3.3:
- Il polinomio $a(x) = x^2 - 2$ è irriducibile in $\mathbb{Q}[x]$. Non lo é peró in $\mathbb{R}[x]$, perché ha $b(x) = x + \sqrt{2}$ come divisore, e puó essere infatti fattorizzato come $x^2 - 2 = \left(x - \sqrt{2}\right)\left(x + \sqrt{2}\right)$;
- Il polinomio $a(x) = x^2 + 1$ è irriducibile in $\mathbb{R}[x]$. Non lo è peró in $\mathbb{C}[x]$, perché ha $b(x) = x + i$ come divisore, e puó essere infatti fattorizzato come $x^2 + 1 = (x - i)(x + i)$.

Teorema 3.3.4 (Teorema fondamentale dell'aritmetica per i polinomi): Per ogni polinomio $p(x) \in K[x]$ tale che $\partial p(x) > 0$ esiste uno ed un solo modo per fattorizzarlo in polinomi primi in $K[x]$ (a meno dell'ordine in cui si dispongono i fattori).

Corollario 3.3.2: Ogni polinomio $a(x) \in K[x]$ di grado $\partial p(x) > 0$ puó essere fattorizzato come $a(x) = ka_1(x)...a_m(x)$, dove $k \in (K - \{0\})$ è il coefficiente direttore di $a(x)$ ed i polinomi $a_1(x), ..., a_m(x)$ sono monici e irriducibili. Tale scrittura è unica, a meno dell'ordine dei fattori.

3.4. Radici di un polinomio

Si consideri un anello di polinomi $K[x]$. Ad un qualsiasi polinomio $f(x) = a_n x^n + ... + a_1 x + a_0$ appartenente a $K[x]$ é possibile associare la funzione $F : K \mapsto K$ cosí definita:

$$F : K \mapsto f(\alpha) = a_n \cdot \alpha^n + ... + a_1 \cdot \alpha + a_0 \ \forall \alpha \in K$$

Siano $f(x) \in K[x]$ e $\alpha \in K$. Se $f(\alpha) = 0$, α si dice **radice** del polinomio $f(x)$.

Teorema 3.4.1 (Teorema di Ruffini): Siano K un campo, $f(x)$ un polinomio in $K[x]$ e α un elemento di K. α è una radice di $f(x)$ se e soltanto se $(x - \alpha)$ é divisore di $f(x)$.

Dimostrazione: Se vale $(x - \alpha) \mid f(x)$, allora si ha $f(x) = (x - \alpha)q(x)$ per un certo $q(x) \in K[x]$. Pertanto, $f(\alpha) = (\alpha - \alpha)q(\alpha) = 0 \cdot q(\alpha) = 0$.
Viceversa, si supponga che α sia una radice di $f(x)$, ovvero che $f(\alpha) = 0$. Per la divisione euclidea fra $f(x)$ e $(x - \alpha)$, esistono due polinomi $p(x)$ e $q(x)$ tali per cui

$$f(x) = (x - \alpha)q(x) + r(x) \text{ con } \partial r(x) < 1$$

Poiché $\partial r(x) < 1$, il polinomio $r(x)$ puó avere esclusivamente grado 0 oppure -1. Nel primo caso, il polinomio é nella forma $r(x) = k$ con $k \in (K - \{0\})$, nel secondo caso il polinomio é il polinomio nullo.
Si noti peró come solamente il secondo caso sia ammissibile. Infatti, se fosse $r(x) = k$ con $k \in (K - \{0\})$, si avrebbe

$$f(x) = (x - \alpha)q(x) + k \Rightarrow f(\alpha) = (\alpha - \alpha)q(\alpha) + k \Rightarrow k = 0$$

Ma questo entra in contraddizione con l'ipotesi che k non sia l'elemento nullo. Pertanto, se ne deduce che $r(x)$ debba per forza essere il polinomio nullo, e che quindi $(x - \alpha)$ divida $f(x)$ senza resto. □

Corollario 3.4.1: Un polinomio $f(x) = ax + b \in K[x]$ di grado 1 (quindi con $a \neq 0$) è irriducibile in $K[x]$ ed ha una ed una sola radice $\alpha \in K$. Tale radice é pari a $-b \cdot a^{-1}$.

Corollario 3.4.2: Sia $f(x)$ un polinomio in $K[x]$ con $\partial f(x) > 1$. Se $f(x)$ ammette radice $\alpha \in K$ allora è riducibile in $K[x]$.

Si noti come il Corollario 3.4.2 non sia una doppia implicazione. Possono infatti esistere dei polinomi $f(x) \in K[x]$ che sono riducibili in $K[x]$ ma che non ammettono radici in K.

Esempio 3.4.1: Il polinomio $x^4 + 3x^2 + 2 \in \mathbb{R}[x]$ si fattorizza come $(x^2 + 1)(x^2 + 2)$ e quindi è riducibile in $\mathbb{R}[x]$, ma non ha radici in $\mathbb{R}$.

Corollario 3.4.3: Un polinomio $f(x) \in K[x]$ di grado 2 oppure 3 è riducibile in $K[x]$ se e solo se ammette una radice in K.

Siano $f(x) \in K[x]$ e $\alpha \in K$. Si dice che α é una radice di $f(x)$ di **molteplicitá algebrica** r, con $r \in \mathbb{N}$ e $r \geq 1$, se $(x - \alpha)^r \mid f(x)$ ma $(x - \alpha)^{r+1} \nmid f(x)$. In particolare, una radice di molteplicità algebrica 1 si dice **radice semplice**.
In altre parole, la molteplicità algebrica di una radice α di un polinomio $p(x) \in K[x]$ indica quante volte il polinomio $(x - \alpha)$ figura come fattore nella fattorizzazione di $p(x)$.

Esempio 3.4.2:
- Il polinomio $f(x) = x^4 - 2x^2 + 1 = (x-1)^2(x+1)^2 \in \mathbb{Q}[x]$ ha in $\mathbb{Q}$ le radici $\alpha_1 = -1$ e $\alpha_2 = 1$ entrambe di molteplicità 2;
- Il polinomio $f(x) = [1]_2 x^4 + [1]_2 = \left([1]_2 x + [1]_2\right)^4 \in \mathbb{Z}_2[x]$ ha in $\mathbb{Z}_2$ la radice $\alpha = [1]_2$ con molteplicità 4.

Teorema 3.4.2: Siano K un campo e $f(x) \in K[x]$ un polinomio non nullo di grado n. La somma delle molteplicità delle radici di $f(x)$ è minore o uguale a n.

Dimostrazione: Se $n = 0$, per definizione $f(x)$ non ha radici in K, pertanto la somma delle molteplicitá delle sue radici é 0. In questo caso il teorema é quindi verificato, perché $n = 0$ e ovviamente $0 \leq 0$.
Se invece $n > 0$, si fattorizzi $f(x)$ in polinomi irriducibili in $K[x]$. Se nessuno di questi ha grado 1, $f(x)$ non ha radici in K. Altrimenti, sia:

$$f(x) = k(x - \alpha_1)^{r_1}(x - \alpha_2)^{r_2}...(x - \alpha_t)^{r_t} g_1(x)...g_m(x)$$

Dove $k \in (K - \{0\})$, $\alpha_1, ..., \alpha_t$ sono elementi distinti di K e $g_1(x), ..., g_m(x)$ sono (se esistono) polinomi di grado maggiore di 1, irriducibili in $K[x]$.
Le radici di $f(x)$ in K sono pertanto $\alpha_1, ..., \alpha_t$ con molteplicità $r_1, ..., r_t$ rispettivamente. É infatti chiaro che α_i è radice con molteplicità r_i, per ciascun $i \in \{1, ..., t\}$. D'altra parte, $f(x)$ non ha altre radici all'infuori di queste, perché se esistesse una radice $\beta \in K$ distinta da ogni α_i si avrebbe:

$$f(\beta) = k(\beta - \alpha_1)^{r_1}...(\beta - \alpha_t)^{r_t} g_1(\beta)...g_m(\beta) \neq 0$$

Che non potrebbe quindi essere una radice.
Confrontando infine il grado di $f(x)$ con il grado di $(x - \alpha_1)^{r_1}...(x - \alpha_t)^{r_t} g_1(x)...g_m(x)$ si trova $r_1 + r_2 + ... + r_t \leq n$. □

4. Crittografia

4.1. Introduzione alla crittografia

Si consideri una situazione in cui si vuole inviare un messaggio opportunamente "occultato" di modo che solamente i destinatari intesi a riceverlo siano in grado di rimuovere l'occultamento e poter leggere il messaggio. Ovvero, non si ha interesse ad impedire a terze parti di poter trovare il messaggio, ma di fare in modo che, anche se terze parti possano intercettarlo, non siano in grado di rimuovere l'occultamento.
Il messaggio originale che si vuole mandare prende il nome di **messaggio in chiaro**, mentre il messaggio opportunamente occultato prende il nome di **messaggio cifrato**. Le due versioni del messaggio sono scritte adoperando i **caratteri** di un certo **alfabeto** (in genere, il medesimo) di dimensione N. Il processo che consiste nel convertire un messaggio in chiaro in un messaggio cifrato prende il nome di **cifratura** o **crittazione**; il processo inverso, ovvero il convertire un messaggio cifrato in un messaggio in chiaro, prende il nome di **decifratura** o **decrittazione**.
Affinché sia possibile manipolarli, il messaggio in chiaro ed il messaggio cifrato devono essere scomposti in elementi atomici, trattabili uno per uno, detti **unitá**. Una unitá puó corrispondere ad un singolo carattere dell'alfabeto su cui i messaggi sono definiti cosí come ad una k-upla di caratteri.
Dovendo poi operare matematicamente su tali unitá per ottenere la cifratura, é necessario tradurre tali caratteri sotto forma di numero. In genere, questo viene fatto determinando un intervallo di numeri interi ed associando ciascuna unitá ad uno di questi numeri interi. Tale associazione deve essere biunivoca, e non é nemmeno necessario nascondere a terze parti la regola che permette tale traduzione.

Esempio 4.1.1: Si voglia spedire il messaggio `A, M, O, G, U, S`. Tale messaggio é scritto nell'alfabeto inglese, avente $N = 26$ caratteri (gli spazi e le virgole non sono parte del messaggio, sono presenti solo per chiarezza).
Si assuma che la conversione fra carattere e numero venga fatta associando a ciascun i-esimo carattere l'i-esimo numero intero nell'intervallo $\{0, \dots, 26\}$, modulo 26. Pertanto, la dimensione delle unitá dei messaggi é pari ad 1. Il messaggio scritto in forma di numero é quindi

$$[1]_{26}, [13]_{26}, [15]_{26}, [7]_{26}, [21]_{26}, [18]_{26}$$

L'operazione di cifratura puó quindi essere vista come una funzione che ha in input una unitá del messaggio in chiaro scritta sotto forma di numero e restituisce una unitá di messaggio cifrato scritta sotto forma di numero. L'operazione di decifratura é la funzione inversa della funzione di cifratura, che ha in input una unitá del messaggio cifrato e restituisce una unitá del messaggio in chiaro. La funzione di cifratura deve essere biunivoca, ovvero ad una unitá di messaggio in chiaro (scritta come numero) deve venire associata una ed una sola unitá di messaggio cifrato (scritta come numero), perché altrimenti non sarebbe possibile costruire la funzione di decifratura (essendo la sua inversa).
Un sistema cosí formulato prende il nome di **sistema crittografico**, e puó essere schematicamente rappresentato come:

$$\mathcal{P} \xrightarrow{f} \mathcal{C} \xrightarrow{f^{-1}} \mathcal{P}$$

- $\mathcal{P}$ é l'insieme di tutte le possibili unitá che costituiscono i messaggi in chiaro, scritte come numero;
- $\mathcal{C}$ é l'insieme di tutte le possibili unitá che costituiscono i messaggi cifrati, scritte come numero;
- f é la funzione di cifratura;
- f^{-1} é la funzione di decifratura (inversa di f).

Esempio 4.1.2:
Si consideri un sistema crittografico cosí costruito (é facile verificare che le funzioni di cifratura e decifratura sono effettivamente l'una l'inversa dell'altra):

$$\mathcal{P} = \mathbb{Z}_N \qquad \mathcal{C} = \mathbb{Z}_N \qquad f(p) = [5]_{26}p + [3]_{26} \qquad f^{-1}(c) = [21]_{26}c - [11]_{26}$$

Si consideri il messaggio del Esempio 4.1.1, scritto come numeri interi modulo 26. La cifratura fornisce:

$$f\left([1]_{26}\right) = [1]_{26} \cdot [5]_{26} + [3]_{26} = [8]_{26} \qquad f\left([13]_{26}\right) = [13]_{26} \cdot [5]_{26} + [3]_{26} = [16]_{26}$$
$$f\left([15]_{26}\right) = [15]_{26} \cdot [5]_{26} + [3]_{26} = [0]_{26} \qquad f\left([7]_{26}\right) = [7]_{26} \cdot [5]_{26} + [3]_{26} = [12]_{26}$$
$$f\left([21]_{26}\right) = [21]_{26} \cdot [5]_{26} + [3]_{26} = [4]_{26} \qquad f\left([18]_{26}\right) = [18]_{26} \cdot [5]_{26} + [3]_{26} = [15]_{26}$$

Mentre la decifratura fornisce:

$$f^{-1}\left([8]_{26}\right) = [21]_{26} \cdot [8]_{26} - [11]_{26} = [1]_{26} \qquad f^{-1}\left([16]_{26}\right) = [21]_{26} \cdot [16]_{26} - [11]_{26} = [13]_{26}$$
$$f^{-1}\left([0]_{26}\right) = [21]_{26} \cdot [0]_{26} - [11]_{26} = [15]_{26} \qquad f^{-1}\left([12]_{26}\right) = [21]_{26} \cdot [12]_{26} - [11]_{26} = [7]_{26}$$
$$f^{-1}\left([4]_{26}\right) = [21]_{26} \cdot [4]_{26} - [11]_{26} = [21]_{26} \qquad f^{-1}\left([15]_{26}\right) = [21]_{26} \cdot [15]_{26} - [11]_{26} = [18]_{26}$$

Effettivamente, i due messaggi coincidono.

In genere, quando ci si riferisce ad un sistema crittografico ci si riferisce ad una *famiglia* di sistemi, che hanno in comune la struttura delle funzioni f e f^{-1} e degli insiemi $\mathcal{P}$ e $\mathcal{C}$ ma differiscono fra di loro per la scelta di determinati parametri. I valori dei parametri della funzione f prendono il nome di **chiave di cifratura**, mentre i valori dei parametri della funzione f^{-1} prendono il nome di **chiave di decifratura**.

Esempio 4.1.3: Il sistema crittografico dell'esempio Esempio 4.1.2 puó essere visto come un membro di una famiglia piú ampia di sistemi crittografici, che hanno questa forma:

$$\mathcal{P} = \mathbb{Z}_N \qquad \mathcal{C} = \mathbb{Z}_N \qquad f(p) = ap + b \qquad f^{-1}(c) = a^{-1}c - a^{-1}b \qquad \text{con } a, b \in \mathbb{Z}_n \text{ e } a \text{ invertibile}$$

Dove tutti i sistemi crittografici di tale famiglia hanno gli stessi insiemi $\mathcal{P}$ e $\mathcal{C}$, ma la scelta delle funzioni di cifratura e di decifratura dipendono dalla coppia di parametri (a, b). Nel caso particolare dell'Esempio 4.1.2, la chiave di cifratura é $\left([5]_{26}, [3]_{26}\right)$, mentre la chiave di decifratura é $\left([21]_{26}, [11]_{26}\right)$.

Di fatto, non é necessario che a terze parti sia nascosto il sistema crittografico che viene usato per cifrare un messaggio, ma é sufficiente che siano le chiavi ad esserlo. Questo perché fintanto che le chiavi di cifratura e di decifratura non sono note, le sole funzioni di cifratura e di decifratura non sono sufficienti a cifrare/decifrare un messaggio.
Essendo la funzione di cifratura e di decifratura l'una l'inversa dell'altra, se la chiave di cifratura viene scoperta, diventa possibile ricostruire la chiave di decifratura, e viceversa. Questo richiede che entrambe le chiavi siano note solo ed esclusivamente alle entitá che si scambiano i messaggi, perché se terze parti riescono ad ottenere l'una possono facilmente determinare l'altra. I sistemi di crittografia che adottano questo approccio prendono il nome di **crittografia a chiave privata**, anche detta **crittografia simmetrica**.

Esempio 4.1.4: Si consideri l'Esempio 4.1.3. Anche se fosse noto il sistema di cifratura, senza almeno una delle due chiavi non é possibile conoscere la vera forma delle due funzioni. Si supponga peró di venire a conoscenza che la chiave di cifratura sia $\left([7]_{26}, [11]_{26}\right)$; diventa allora possibile determinare la chiave di decifratura (e quindi la vera forma della funzione di decifratura) con pochi passi:

$$f^{-1}(c) = a^{-1}c - a^{-1}b \Rightarrow f^{-1}(c) = \left([7]_{26}\right)^{-1} c - \left([7]_{26}\right)^{-1}[11]_{26} \Rightarrow f^{-1}(c) = [15]_{26}c - [9]_{26}$$

Nella **crittografia a chiave pubblica**, anche detta **crittografia asimmetrica**, é invece possibile rendere nota la chiave di cifratura senza che questo comporti che si possa usarla per ricavare la chiave di decifratura. Questo perché la funzione di cifratura viene appositamente scelta di modo che, anche ammesso di conoscere la chiave di cifratura, calcolare la chiave di decifratura a partire da questa richieda una computazione troppo lunga per essere ragionevole.
Funzioni per le quali é semplice valutarle nel loro input ma che é proibitivo calcolarne l'inversa sono dette **one-way function**. La nozione di one-way function non é rigorosa dal punto di vista matematico, dato che il tempo necessario per calcolare la funzione inversa di una funzione dipende anche dalla tecnologia attualmente a disposizione. In parole povere, un incremento nella potenza dei calcolatori puó rendere funzioni un tempo considerate one-way function delle funzioni "comuni".
Sarebbe tecnicamente possibile provare matematicamente che il calcolo dell'inversa di una certa funzione sia un problema non risolvibile in tempo polinomiale. In questo modo, non importa la potenza dei calcolatori, tale computazione sará sempre e comunque improponibile (se non per piccole istanze). É peró interessante notare come non esista alcuna prova matematica che il calcolo di funzioni inverse per funzioni di cifratura comunemente utilizzabili sia un problema intrattabile.

4.2. Algoritmo RSA

RSA é un esempio di sistema crittografico asimmetrico. Siano Alice e Bob due entitá che hanno intenzione di comunicare scambiandosi messaggi senza che terze parti possano conoscerne il contenuto (ovvero, anche ammesso che possano intercettare il messaggio, non possano decifrarlo). Si assuma che Alice sia il ricevente e Bob il mittente. La cifratura e decifratura di messaggi mediante RSA puó essere descritta sotto forma di algoritmo:

1. Alice sceglie una coppia di numeri primi distinti, siano questi p e q;
2. Alice calcola il loro prodotto $N = pq$ ed il valore di $\varphi(N)$, che per le proprietá di tale funzione é semplicemente $(p-1)(q-1)$;
3. Alice sceglie un numero casuale r tale che sia coprimo con $\varphi(N)$ e piú piccolo di quest'ultimo;
4. Alice calcola l'identitá di Bézout per r e $\varphi(N)$, ovvero determina una coppia di numeri interi s e t tali per cui $rs + \varphi(N)t = 1$;
5. Alice rende pubblica la chiave di cifratura (N, r), mentre tiene per sé i numeri p, q e $\varphi(N)$, cosí come la chiave di decifratura (N, s);
6. Sia il messaggio che Bob vuole mandare ad Alice il numero intero b compreso fra 0 ed N. Bob legge la chiave di cifratura di Alice ed invia ad Alice il numero $a = b^r \bmod N$;
7. Alice riceve a e ricostruisce il messaggio b originale come $b = a^s \bmod N$;

Teorema 4.2.1 (Correttezza dell'algoritmo RSA): L'algoritmo RSA é corretto. Ovvero, il messaggio decifrato da Alice coincide sempre con il messaggio inviato da Bob.

Dimostrazione: Si consideri il caso in cui b ed N siano coprimi. Dovendo esistere s e t tali per cui $rs + \varphi(N)t = 1$, si ha:

$$b = b^1 \bmod N = b^{rs+\varphi(N)t} \bmod N = (b^{rs})\left(b^{\varphi(N)t}\right) \bmod N = \left((b^r)^s \bmod N\right)\left(\left(b^{\varphi(N)}\right)^t \bmod N\right)$$

Per il Teorema 2.9.2, si ha $b^{\varphi(N)} \equiv 1 \bmod N$, in quanto b e N sono stati assunti coprimi per ipotesi, ed a maggior ragione $\left(b^{\varphi(N)}\right)^t \equiv 1 \bmod N$. Pertanto:

$$((b^r)^s \bmod N)\left(\left(b^{\varphi(N)}\right)^t \bmod N\right) = ((b^r)^s \bmod N)(1 \bmod N) = a^s \bmod N$$

Si consideri invece il caso in cui b ed N non siano coprimi. Essendo N il prodotto di due numeri primi distinti p e q ed avendo scelto come inferiore ad N, b deve essere multiplo o di p o di q. Si assuma, senza perdita di generalitá, che b sia multiplo di p, ovvero che esista un $q \in \mathbb{Z}$ maggiore di k tale per cui $b = kp$. Essendo p e q primi ed avendo assunto che b sia multiplo di q, deve aversi che q e b siano coprimi. Per il Teorema 2.9.2, deve valere $b^{\varphi(q)} \equiv 1 \bmod q$, dove $\varphi(q) = q - 1$ per il Lemma 2.8.1. Si ha poi:

$$b^{\varphi(N)} = b^{\varphi(pq)} = b^{(q-1)(p-1)} = \left(b^{q-1}\right)^{p-1} \equiv b^{-t\varphi(N)} \equiv 1 \bmod q$$

La congruenza $b^{-t\varphi(N)} \equiv 1 \bmod q$ equivale a $b^{-t\varphi(N)} = 1 + wq$ per un certo $w \in \mathbb{Z}$. Si ha:

$$b^{-t\varphi(N)} = 1 + wq \Rightarrow b^{1-t\varphi(N)} = b + bwq \Rightarrow b^{1-t\varphi(N)} = b + wkN \Rightarrow b^{rs} = b + wkN \equiv b \bmod N$$

Ovvero, anche in questo caso:

$$a^s \bmod N = (b^r)^s \bmod N = b^{rs} \bmod N = b$$

□

Esempio 4.2.1: Si supponga che Bob voglia inviare ad Alice un messaggio, occultandolo usando il sistema crittografico RSA. Si assuma innanzitutto che entrambi sappiano che la conversione carattere/numero venga fatta con unitá di grandezza 1 mediante questo schema:

A	B	C	D	E	F	G	H	I	L	M	N	O	P	Q	R	S	T	U	V	Z	∎
2	3	4	5	6	7	8	9	29	31	12	13	14	37	16	17	18	19	43	21	22	23

Alice sceglie i numeri primi $p = 5$ e $q = 11$. A partire da questi, Alice calcola:

$$N = pq = 55 \quad \varphi(N) = \varphi(55) = \varphi(5) \cdot \varphi(11) = (5-1) \cdot (11-1) = 40$$

Alice sceglie poi il numero $r = 37$ tale che $r < \varphi(N)$ e $\mathrm{MCD}(r, \varphi(N)) = \mathrm{MCD}(37, 40) = 1$. A partire da questi, Alice calcola due interi s e t per i quali $37s + 40t = 1$, mediante l'algoritmo di Euclide:

$$\begin{aligned} 40 &= 37 \cdot 1 + 3 \\ 37 &= 3 \cdot 12 + 1 \\ 3 &= 3 \cdot 1 + 0 \end{aligned} \qquad \begin{aligned} 3 &= 40 - 37 \\ 1 &= 37 - 12 \cdot 3 = 37 - 12(40 - 37) \\ &= -12 \cdot 40 + 13 \cdot 37 \end{aligned}$$

Ottenendo $t = -12$ e $s = 13$. A questo punto, Alice ha ricavato la chiave di cifratura $(N, r) = (55, 37)$, e la rende pubblica.
Bob legge le informazioni rese pubbliche da Alice e le spedisce il messaggio seguente:

26	7	21	9	52	7	52	41	23	28	24	7	18	49	7

Alice decodifica ciascuna unitá del messaggio a parire dall'equivalenza $b = a^s \bmod N$:

$$\begin{gathered} 26^{13} \bmod 55 = 31 \quad 7^{13} \bmod 55 = 2 \quad 21^{13} \bmod 55 = 21 \quad 9^{13} \bmod 55 = 14 \quad 52^{13} \bmod 55 = 17 \\ 7^{13} \bmod 55 = 2 \quad 52^{13} \bmod 55 = 17 \quad 41^{13} \bmod 55 = 6 \quad 23^{13} \bmod 55 = 23 \quad 28^{13} \bmod 55 = 18 \\ 24^{13} \bmod 55 = 19 \quad 7^{13} \bmod 55 = 2 \quad 18^{13} \bmod 55 = 13 \quad 49^{13} \bmod 55 = 4 \quad 7^{13} \bmod 55 = 2 \end{gathered}$$

Il messaggio decodificato é quindi:

31	2	21	14	17	2	17	6	23	18	19	2	13	4	2

Convertendo ordinatamente ciascuna unitá da numero a carattere, si ottiene:

L	A	V	O	R	A	R	E	∎	S	T	A	N	C	A

La funzione di cifratura dell'algoritmo RSA é effettivamente una one-way function perché per conoscere il termine s della chiave di decifratura é necessario risolvere $rs + \varphi(N)t = 1$, che a sua volta richiede di calcolare $\varphi(N)$. Il problema é che, per le proprietá della funzione di Eulero, tale valore é facile da calcolare solamente se é nota la fattorizzazione in numeri primi di N, e tale informazione é nota solamente ad Alice. Sebbene sarebbe tecnicamente possibile determinare la fattorizzazione in numeri primi di qualsiasi intero, realisticamente questo richiede tempi troppo lunghi, specialmente se l'intero in questione é molto grande[1].

4.3. Firma digitale tramite RSA

Nello scenario presentato nel precedente capitolo, due entitá sono in grado di scambiarsi messaggi senza che terze parti siano in grado di leggerne il contenuto. Si presenta peró un problema: per il ricevente non c'é modo di sapere *davvero* se il mittente del messaggio sia effettivamente chi dice di essere o se sia qualcun'altro che lo sta impersonando. Diventa quindi necessario associare al messaggio una **firma** che dia al ricevente la certezza di aver ricevuto un messaggio da un mittente che é davvero chi dice di essere. Una firma di questo tipo puó essere facilmente implementata mediante RSA.
Siano Alice e Bob due entitá che vogliono comunicare; senza perdita di generalitá, si assuma che Alice sia il mittente e Bob il ricevente. Alice ha una chiave di cifratura (N_A, r_A) ed una chiave di decifratura (N_A, s_A), mentre Bob ha una chiave di cifratura (N_B, r_B) ed una chiave di decifratura (N_B, s_B). Naturalmente, le chiavi di cifratura sono note per entrambi, mentre le chiavi di decifratura sono note solamente ai rispettivi possessori.
Si assuma per semplicitá che $\mathcal{P}$, l'insieme delle unitá di messaggi in chiaro, e $\mathcal{C}$, l'insieme delle unitá dei messaggi cifrati, coincidano. Sia F un messaggio non cifrato speciale (un numero identificativo, un timestamp, ecc...) che fa da firma ai messaggi di Alice. Si noti come Alice, per provare a Bob di essere sé stessa e non un impostore, non puó semplicemente cifrare ed inviare F, perché si ripresenterebbe il problema. Si distinguono due casi:

1. $N_A \geq N_B$. Per cifrare il suo messaggio, Alice innanzitutto cifra F usando la sua chiave di decifratura (anziché quella di cifratura, come farebbe di norma), ottenendo $F_A = F^{s_A} \bmod N_A$. A partire da questa, calcola $F_{A,B} = F_A^{r_B} \bmod N_B$ e invia a Bob sia il messaggio che vuole inviargli, sia la sua firma $F_{A,B}$. Bob ricostruisce sia il messaggio sia la firma $F_A = F_{A,B}^{s_B} N_B$, per poi ottenere F come $F = F_A^{r_A} \bmod N_A$;
2. $N_A < N_B$. Per cifrare il suo messaggio, Alice innanzitutto cifra F usando la chiave di cifratura di Bob (anziché la sua chiave di cifratura, come farebbe di norma), ottenendo $F_B = F^{r_B} \bmod N_B$. A partire da questa, calcola $F_{B,A} = F_B^{s_A} \bmod N_A$ e invia a Bob sia il messaggio che vuole inviargli, sia la sua firma $F_{B,A}$. Bob ricostruisce sia il messaggio sia la firma $F_B = F_{B,A}^{r_A} N_A$ per poi ottenere F come $F = F_B^{s_B} \bmod N_B$.

Bob ha la certezza che il messaggio provenga da Alice perché per ricostruire la firma ha dovuto usare la chiave pubblica di Alice, e solamente Alice puó conoscere la relativa chiave privata, e aver quindi cifrato il messaggio con essa.

[1] A dire il vero, non é mai stato dimostrato che non possa esistere un algoritmo in grado di calcolare velocemente la fattorizzazione in numeri primi. Per tale motivo, al momento questa é soltanto una congettura.

5. Teoria dei codici

5.1. Introduzione alla teoria dei codici

Prende il nome di **sistema di comunicazione** la struttura di seguito schematizzata:

$$\begin{array}{ccccccccc} & & & & \text{rumore} & & & & \\ & & & & \downarrow & & & & \\ \text{Trasmettitore} & \xrightarrow{m} & \text{Codificatore} & \xrightarrow{x} & \text{Canale} & \xrightarrow{y} & \text{Decodificatore} & \xrightarrow{\widetilde{m}} & \text{Ricevitore} \end{array}$$

La descrizione dei componenti é qui riportata:

- Trasmettitore: emette il messaggio m;
- Codificatore: traduce il messaggio m nella parola x in modo che possa attraversare il canale;
- Canale: mezzo attraverso il quale viaggiano le parole;
- Decodificatore: trasforma la parola y in uscita dal canale nel messaggio $\widetilde{m}$;
- Ricevitore: riceve il messaggio $\widetilde{m}$;
- Rumore: disturbi di vario genere che potrebbero alterare le parole.

In una situazione ideale, il segnale inviato x ed il segnale ricevuto y dovrebbero coincidere. In uno scenario piú realistico, i due segnali saranno piú o meno diversi, in quanto ogni canale di comunicazione é soggetto a rumore, e quindi parte dell'informazione giunta a destinazione differirá dall'originale. In termini molto generali, i tipi di errori che possono presentarsi nella trasmissione di un segnale x e nella ricezione del segnale y sono tre:

- Parte dell'informazione contenuta in x viene alterata;
- Parte dell'informazione contenuta in x viene perduta;
- Il segnale x si ritrova ad avere piú informazioni dell'originale quando viene ricevuto.

Essendo la presenza di tali errori inevitabile, l'interesse é quello di costruire canali di comunicazione che, pur essendo vulnerabili al rumore, sono comunque in grado di tollerarlo, di modo che il messaggio ricevuto $\widetilde{m}$ sia una buona approssimazione di quello inviato m.

5.2. Codici a blocchi

Sia $A_q = \{x_1, x_2, ..., x_q\}$ un insieme finito di cardinalità q, con $q \geq 2$. Prende il nome di **codice a blocchi** un qualunque sottoinsieme non vuoto C di A_q^n. In particolare:

- A_q viene detto *alfabeto* di C;
- A_q^n viene detto *spazio delle parole* di lunghezza n (nell'alfabeto A_q);
- Una **parola** del codice C è una qualsiasi n-upla ordinata di simboli dell'alfabeto A_q;
- n viene anche chiamata *lunghezza* del codice;
- La cardinalità di C viene chiamata *grandezza* del codice.

La notazione matematica per le n-uple ordinate sarebbe $(x_1, x_2, ..., x_n)$, ma per semplicitá verranno omesse sia le parentesi, sia le virgole. Inoltre, spesso si ha $A_q = \{0, 1\}$, perché l'interesse della trasmissione dei segnali é prettamente informatico.

Esempio 5.2.1: Si consideri l'alfabeto $A_q = \{0, 1\}$. Il codice C sull'alfabeto A_q riportato di seguito é di lunghezza 3 e di grandezza (numero di parole) 3:

$$C = \{001, 010, 100\} \qquad A_q^3 = A_q \times A_q \times A_q = \{000, 001, 010, 011, 100, 101, 110, 111\}$$

Si supponga di aver inviato una parola $p = (x_1, ..., x_n)$ e di aver ricevuto una parola $p' = (y_1, ..., y_n)$. Se queste differiscono, allora significa che si é in presenza di un errore. Per semplicitá, si considerino solamente errori di primo tipo, ovvero che uno o piú simboli di p non corrispondano ai rispettivi simboli in p'. Il numero di errori verrá conteggiato in base al numero di coppie di simboli che differiscono (una coppia di simboli diversi é un errore, due coppie di simboli diversi sono due errori, ecc...). Si assuma inoltre che gli errori siano *eventi indipendenti*, ovvero che se $x_i \neq y_i$ per una certa posizione i questo non influenza il verificarsi di un errore in un'altra posizione $j \neq i$.

Per misurare quanto p e p' sono "dissimili", é necessario introdurre una misura di *distanza*. La forma di distanza maggiormente utilizzata in questo contesto é la **distanza di Hamming**:

$$d : A_q^n \times A_q^n \mapsto \mathbb{R},\ d(p, p') = |\{i : x_i \neq y_i\}|$$

Ovvero, la distanza di Hamming é pari al numero di simboli delle due parole nella stessa posizione che differiscono (a prescindere da quali siano i simboli in questione). Dato che in questo contesto verrá sempre usata la distanza di Hamming come forma di distanza, si sottointenderá con il solo termine "distanza" la distanza di Hamming.
Essendo una distanza, la distanza di Hamming gode, per qualsiasi parola sull'alfabeto A_q, delle seguenti quattro proprietá:

1. $d(p, p') = d(p', p)$;
2. $d(p, p') = 0$ se e soltanto se $p = p'$;
3. $d(p, p') \geq 0$;
4. É verificata la **disuguaglianza triangolare**, ovvero $d(p, p') \leq d(p, p'') + d(p'', p')$.

Dato un codice $C \subseteq A_q^n$, si dice **distanza minima** di C il minimo delle distanze tra due parole distinte di C:

$$d(C) = \min\{d(p, p') : p, p' \in C, p \neq p'\}$$

Esempio 5.2.2: Sia $A_2 = \{0, 1\}$ un alfabeto. Sia poi $C = \{000, 001, 010, 100, 111\}$ un codice su A_2 di lunghezza 3. Le distanze fra ciascuna coppia di parole di C, escludendo le coppie ripetute e le distanze fra ciascuna parola e sé stessa, sono:

$$d(000, 001) = 1 \quad d(000, 010) = 1 \quad d(000, 100) = 1 \quad d(000, 111) = 3 \quad d(001, 010) = 2$$
$$d(001, 100) = 2 \quad d(001, 111) = 2 \quad d(010, 100) = 2 \quad d(010, 111) = 2 \quad d(100, 111) = 2$$

Pertanto, $d(C) = 1$.

Si assuma che due entitá abbiano a disposizione il medesimo codice, e che l'una invii all'altra una parola. Si supponga che tale parola ricevuta non sia presente nel codice; se ne deduce che questa sia stata danneggiata durante la trasmissione. La parola che é ragionevole assumere sia stata inviata in origine é quella presente nel codice che maggiormente somiglia a quella ricevuta, fintanto che la differenza fra le due é sufficientemente piccola. Tale principio viene detto **principio di massima verosimiglianza**.
Si supponga di avere a disposizione un codice C e di aver ricevuto la parola $w \in A_q^n$. Il codice C **corregge** la parola w se e soltanto se esiste una ed una sola parola in C a distanza minima da w, cioè se e soltanto se esiste una ed una sola $x \in C$ tale per cui $d(x, w) = \min\{d(y, w) : y \in C\}$. In tal caso, w viene corretta con x.
Si noti come non sia garantito che la correzione del codice restituisca la parola che é stata effettivamente inviata. Infatti, la parola ricevuta viene corretta (se necessario) con la parola che meno dista da questa, ma se il canale di comunicazione é particolarmente rumoroso la parola ricevuta potrebbe differire profondamente da quella originale, e molto piú simile ad una parola diversa. Inoltre, non é nemmeno garantito che la parola del codice con minima distanza da quella ricevuta sia sempre una sola, perché potrebbero esistere piú parole con la medesima distanza minima.

Esempio 5.2.3: Sia $C = \{000000, 111111, 222222\}$ un codice di lunghezza 6 sull'alfabeto $A_3 = \{0, 1, 2\}$. Si supponga che Alice invii a Bob la parola 000000, e che Bob corregga la parola ricevuta impiegando il principio di massima verosimiglianza.

- Si supponga che Bob riceva la parola 001102. Poiché:

  $$d(000000, 001102) = 3, d(111111, 001102) = 4, d(222222, 001102) = 5$$

  Bob corregge (correttamente) la parola ricevuta con 000000.
- Si supponga che Bob riceva la parola 022220. Poiché:

  $$d(000000, 022220) = 4, d(111111, 022220) = 6, d(222222, 022220) = 2$$

  Bob corregge (erroneamente) la parola ricevuta con 222222.
- Si supponga che Bob riceva la parola 000111. Poiché:

  $$d(000000, 000111) = 3, d(111111, 000111) = 3, d(222222, 000111) = 6$$

  Bob non é in grado di correggere la parola ricevuta, perché esistono piú parole con la stessa distanza da questa.

Un codice $C \subseteq A_q^n$ si dice k-**rivelatore** se k è il numero massimo di errori che è in grado di rivelare.

Teorema 5.2.1: Un codice $C \subseteq A_q^n$ è k-rivelatore se e soltanto se $d(C) = k + 1$.

Dimostrazione: Si supponga che $d(C) = k + 1$. Sia p la parola inviata, e sia p' la parola ricevuta. Sia poi t il numero di errori subiti da p durante la trasmissione, ovvero $d(p, p') = t$. Si distinguono due casi:

- $t < k$. Allora $d(p, p') = t < k < k + 1 = d(C) = \min\{d(w, w') : w, w' \in C, w \neq w'\}$. Questo significa che $p' \notin C$, e che quindi i t errori vengono rivelati. Essendo $t < k$, a maggior ragione i k errori verranno tutti rivelati;
- $t \geq k$. Allora $d(p, p') = t \geq k = d(C) - 1 = \min\{d(w, w') : w, w' \in C, w \neq w'\} - 1$. Questo significa che potrebbe aversi $p' \in C$, e che quindi possa esistere un errore fra i t che non viene rivelato. Essendo $t \geq k$, non vi é garanzia che tutti i k errori verranno rivelati.

Si ha quindi che C é k-rivelatore. Viceversa, sia k il massimo numero di errori che C é in grado di rivelare. Ogni parola $p'' \in C$ distinta da p deve differire da questa in almeno $k + 1$ componenti, pertanto si ha $d(C) \geq k + 1$. Inoltre, poichè C rivela k errori ma non $k + 1$, devono esistere due parole $w, w' \in C$ tali per cui $d(w, w') = k + 1$. Ne consegue che $d(C) = k + 1$. □

Corollario 5.2.1: Un codice $C \subseteq A_q^n$ rivela t errori se e solo se $d(C) \geq t + 1$.

Dimostrazione: Il codice C rivela t errori se e solo se alterando una parola di C in $r \leq t$ componenti non si ottiene un'altra parola di C. Questo avviene se e solo se due parole di C distano almeno $t + 1$. □

Un codice $C \subseteq A_q^n$ si dice h-**correttore** se h è il numero massimo di errori che è in grado di correggere.

Teorema 5.2.2: Ogni codice $C \subseteq A_q^n$ é $\left\lfloor \frac{d(C)-1}{2} \right\rfloor$-correttore.

Dimostrazione: Siano $p, p' \in C$ rispettivamente la parola trasmessa e la parola ricevuta, con t numero di errori subiti da p durante la trasmissione. Si supponga poi $t \leq \left\lfloor \frac{d(C)-1}{2} \right\rfloor$. Affinché C sia $\left\lfloor \frac{d(C)-1}{2} \right\rfloor$-corret-

tore, la parola p che viene scelta come correzione per p' deve essere l'unica e sola parola in C che dista da p' meno di tutte. In altre parole, qualsiasi parola p'' distinta da p dev'essere piú distante da p' di quanto p' disti da p. Formalmente:

$$\forall p'' \in C, p'' \neq p \text{ si ha } d(p'', p') > d(p, p') = t$$

Avendo supposto $t \leq \left\lfloor \frac{d(C)-1}{2} \right\rfloor$, questo equivale a dimostrare che:

$$\forall p'' \in C, p'' \neq p \text{ si ha } d(p'', p') > \left\lfloor \frac{d(C)-1}{2} \right\rfloor$$

Si supponga per assurdo che questo non sia vero, e che esista quindi una parola $p''' \in C$ distinta da p tale per cui $d(p''', p') \leq \left\lfloor \frac{d(C)-1}{2} \right\rfloor$. Applicando la disuguaglianza triangolare, si ha:

$$d(p''', p) \leq d(p''', p') + d(p', p) \leq \left\lfloor \frac{d(C)-1}{2} \right\rfloor + \left\lfloor \frac{d(C)-1}{2} \right\rfloor = 2\left\lfloor \frac{d(C)-1}{2} \right\rfloor$$

Per definizione di arrotondamento per difetto, $\lfloor a \rfloor = a - \varepsilon$ con $\varepsilon \in \mathbb{R}$ tale che $0 \leq \varepsilon < 1$. Si ha quindi:

$$d(p''', p) \leq 2\left(\frac{d(C)-1}{2} - \varepsilon\right) = \frac{\cancel{2}(d(C)-1)}{\cancel{2}} - 2\varepsilon = d(C) - 1 - 2\varepsilon \leq d(C)$$

Questo peró non é possibile, perché per ipotesi $d(C)$ é la minima distanza fra due parole in C. Pertanto, occorre assumere che p''' non possa esistere. □

Corollario 5.2.2: Un codice $C \subseteq A_q^n$ corregge t errori se e solo se $d(C) \geq 2t + 1$.

5.3. Codici lineari

Sia $\mathbb{Z}_p^n$ lo spazio vettoriale delle n-uple di $\mathbb{Z}_p$, con p numero primo. La somma tra due vettori su tale spazio vettoriale é definita come:

$$\left([x_1]_p, ..., [x_n]_p\right) + \left([y_1]_p, ..., [y_n]_p\right) = \left([x_1 + y_1]_p, ..., [x_n + y_n]_p\right)$$

Mentre il prodotto fra un vettore ed uno scalare come:

$$[\lambda]_p\left([x_1]_p, ..., [x_n]_p\right) = \left([\lambda x_1]_p, ..., [\lambda x_n]_p\right)$$

La base canonica di tale spazio vettoriale viene definita come:

$$\left([1]_p, [0]_p, ..., [0]_p, [0]_p\right), \left([0]_p, [1]_p, ..., [0]_p, [0]_p\right), ..., \left([0]_p, [0]_p, ..., [1]_p, [0]_p\right), \left([0]_p, [0]_p, ..., [0]_p, [1]_p\right)$$

Per semplicitá, quando é desumibile dal contesto, una classe di resto $[x]_p$ verrá semplicemente denotata con x.

Esempio 5.3.1: Lo spazio vettoriale $\mathbb{Z}_2^5$, é costituito da tutte e sole le quintuple che hanno per elementi gli elementi di $\mathbb{Z}_2 = \{0, 1\}$ (si noti come con 0 si intenda $[0]_2$, mentre con 1 si intenda $[1]_2$).

Teorema 5.3.1: Lo spazio vettoriale $\mathbb{Z}_p^n$ ha dimensione n.

Un qualsiasi sottospazio vettoriale di $\mathbb{Z}_p^n$ viene detto **codice lineare**.

Esempio 5.3.2: A partire dallo spazio vettoriale $\mathbb{Z}_2^5$ é possibile definire il codice C come il suo sottospazio avente base $B = \{b_1 = 10111, b_2 = 11110\}$. I vettori che costituiscono C sono tutti e i soli vettori generati dalla combinazione lineare $\lambda_1 b_1 + \lambda_2 b_2$, con $\lambda_1, \lambda_2 \in \mathbb{Z}_2$. Essendo $\mathbb{Z}_2$ e B due insiemi finiti, é possibile enumerare C esplicitamente:

$$0(10111) + 0(11110) = (00000) \qquad 1(10111) + 0(11110) = (10111)$$
$$0(10111) + 1(11110) = (11110) \qquad 1(10111) + 1(11110) = (01001)$$

Lemma 5.3.1: Sia $C \subseteq \mathbb{Z}_p^n$ un codice lineare di dimensione k. Si ha $|C| = p^k$.

Dimostrazione: Sia $B = \{b_1, b_2, ..., b_k\}$ una base di C. Ogni elemento di C puó essere generato a partire da una ed una sola combinazione lineare dei vettori di B, ovvero

$$\lambda_1 b_1 + \lambda_2 b_2 + \cdots + \lambda_k b_k \quad \text{con} \quad \lambda_i \in \mathbb{Z}_p, \ i = \{1, ..., k\}$$

Essendo i λ_i esattamente p e dovendone scegliere k per generare ciascun vettore, anche ripetuti, il numero totale di vettori di C é p^k. □

Lemma 5.3.2: Ogni codice lineare C contiene[2] la parola $\underline{0} = 00...0$.

Dimostrazione: Sia $B = \{b_1, b_2, ..., b_k\}$ una base di C e sia k la sua dimensione. Per generare $00...0$ occorre costruire una combinazione lineare dove tutti gli elementi sono nulli. Questo é sempre possibile perché, per qualsiasi p, l'elemento $[0]_p$ appartiene a $\mathbb{Z}_p$, ed quindi é sempre possibile costruire una combinazione lineare del tipo:

$$0b_1 + 0b_2 + \cdots + 0b_k \quad \text{ovvero} \quad \lambda_1 = \lambda_2 = ... = \lambda_k = 0$$

□

Sia $x = (x_1 x_2 ... x_n)$ un elemento di $\mathbb{Z}_p^n$. Prende il nome di **peso di Hamming** il numero di componenti $w(x)$ di x diverse da 0, ovvero:

$$w(x) = \{i \mid x_i \neq 0\}$$

Dato che in questo contesto verrá sempre usato il peso di Hamming come nozione di peso, si sottointenderá con il solo termine "peso" il peso di Hamming.

Lemma 5.3.3: Sia C un codice lineare, e siano x, y due suoi elementi. Allora $d(x, y) = w(x - y)$.

Dimostrazione: Per il Lemma 5.3.2, la parola $\underline{0}$ appartiene sempre a C. Pertanto, per qualsiasi $y \in C$, vale $d(\underline{0}, y) = w(y)$, perché di fatto le due definizioni coincidono. Poichè $d(C)$ è la distanza minima di C esistono certamente $x, y \in C$ con $d(C) = d(x, y) = w(x - y)$. □

[2] É inoltre vero per definizione, dato che $\underline{0}$ é il vettore nullo di $\mathbb{Z}_p^n$ e qualsiasi sottospazio vettoriale deve contenerlo.

Lemma 5.3.4: Sia C un codice lineare. La distanza minima $d(C)$ è pari al peso della parola non nulla di C avente, fra tutte, il peso minimo. Ovvero:

$$d(C) = \min\{w(z) : z \in C, z \neq \underline{0}\}$$

Dimostrazione: Si supponga per assurdo che $d(C) < \min\{w(z) : z \in C, z \neq \underline{0}\}$. Se cosí fosse, potrebbe esisterebbe un $z_0 \in C$ tale per cui $w(z_0) = d(z_0, \underline{0}) < d(C)$, ma questo non é possibile, perché $d(C)$ é la minima distanza fra tutte le parole di C. □

Sia $C \in \mathbb{Z}_p^n$ un codice lineare di dimensione k. Siano poi $\mathcal{B}_C = \{b_1, b_2, ..., b_k\}$ una base di C e $\mathcal{B} = \{e_1, e_2, ..., e_k\}$ una base di $\mathbb{Z}_p^n$. Ciascun vettore $b_i \in \mathcal{B}_C$ puó essere scritto come combinazione lineare a coefficienti $\mathbb{Z}_p$ dei vettori di $\mathcal{B}$:

$$\begin{cases} b_1 = \lambda_{1,1}e_1 + \lambda_{1,2}e_2 + ... + \lambda_{1,n}e_n \\ b_2 = \lambda_{2,1}e_1 + \lambda_{2,2}e_2 + ... + \lambda_{2,n}e_n \\ \vdots \\ b_k = \lambda_{k,1}e_1 + \lambda_{k,2}e_2 + ... + \lambda_{k,n}e_n \end{cases} \text{ con } \lambda_{i,j} \in \mathbb{Z}_p \ \forall i, j = \{1, ..., k\}$$

La matrice G costituita dai coefficienti $\lambda_{i,j}$ di tali combinazioni lineari, ovvero:

$$G = \begin{pmatrix} \lambda_{1,1} & \lambda_{1,2} & \dots & \lambda_{1,n} \\ \lambda_{2,1} & \lambda_{2,2} & \dots & \lambda_{2,n} \\ \vdots & \dots & \ddots & \vdots \\ \lambda_{k,1} & \lambda_{k,2} & \dots & \lambda_{k,n} \end{pmatrix} \in \mathrm{Mat}\big(k \times n, \mathbb{Z}_p\big)$$

Viene detta **matrice generatrice** di G.
La matrice G puó essere usata per la codifica dei messaggi, ovvero per associare ad un vettore $m \in Z_p^k$ una parola in $C \subseteq Z_p^k$. Dato un vettore $m = (m_1, m_2, ..., m_k) \in Z_p^k$, la codifica di m è rispetto a G é data dal prodotto matriciale fra m e G, ovvero:

$$mG = (m_1 \ \ m_2 \ \ \dots \ \ m_k) \begin{pmatrix} \lambda_{1,1} & \lambda_{1,2} & \dots & \lambda_{1,n} \\ \lambda_{2,1} & \lambda_{2,2} & \dots & \lambda_{2,n} \\ \vdots & \dots & \ddots & \vdots \\ \lambda_{k,1} & \lambda_{k,2} & \dots & \lambda_{k,n} \end{pmatrix} \in C$$

Esempio 5.3.3: Si consideri il codice lineare $C \subseteq \mathbb{Z}_2^5$ di dimensione 3. Sia $\mathcal{B}_C$ la base di C costituita dai vettori $\{b_1 = 10001, b_2 = 11010, b_3 = 11101\}$. Sia poi $\mathcal{B}$ la base canonica di $\mathbb{Z}_2^5$.
La matrice G é cosí costruita:

$$\begin{cases} 10001 = \lambda_{1,1}10000 + \lambda_{1,2}01000 + \lambda_{1,3}00100 + \lambda_{1,4}00010 + \lambda_{1,5}00001 \\ 11010 = \lambda_{2,1}10000 + \lambda_{2,2}01000 + \lambda_{2,3}00100 + \lambda_{2,4}00010 + \lambda_{2,5}00001 \\ 11101 = \lambda_{3,1}10000 + \lambda_{3,2}01000 + \lambda_{3,3}00100 + \lambda_{3,4}00010 + \lambda_{3,5}00001 \end{cases} \Rightarrow G = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \end{pmatrix}$$

Sia infine $m = (1, 0, 1) \in \mathbb{Z}_2^3$ il messaggio da codificare. La codifica di m rispetto a C é:

$$mG = (1 \ \ 0 \ \ 1) \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 \cdot 1 + 0 \cdot 1 + 1 \cdot 1 \\ 1 \cdot 0 + 0 \cdot 1 + 1 \cdot 1 \\ 1 \cdot 0 + 0 \cdot 0 + 1 \cdot 1 \\ 1 \cdot 0 + 0 \cdot 1 + 1 \cdot 0 \\ 1 \cdot 1 + 0 \cdot 0 + 1 \cdot 1 \end{pmatrix} = \begin{pmatrix} 1 + 0 + 1 \\ 0 + 0 + 1 \\ 0 + 0 + 1 \\ 0 + 0 + 0 \\ 1 + 0 + 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \in C$$