Chapter 3

Interi e aritmetica modulare

In questo capitolo lavoreremo con i numeri interi: ricorderemo i fatti fondamentali dell'aritmetica che ci serviranno nelle sezioni successive per introdurre la cosiddetta aritmetica modulare, che è di fondamentale importanza in molte applicazioni pratiche.

3.1 Somma, prodotto e algoritmo della divisione

Iniziamo con il ricordare le proprietà fondamentali delle operazioni di somma e prodotto sull'insieme $\mathbb{Z}$ dei numeri interi:

(1) *la somma è associativa*, ovvero per ogni $a, b, c \in \mathbb{Z}$ si ha

$$(a + b) + c = a + (b + c)$$

(2) *la somma è commutativa*, ovvero per ogni $a, b \in \mathbb{Z}$ si ha

$$a + b = b + a$$

(3) *esiste un elemento neutro per la somma* (lo 0), tale che per ogni $a \in \mathbb{Z}$ si ha

$$a + 0 = 0 + a = a$$

(4) *ogni elemento ammette un inverso rispetto alla somma* (che chiameremo anche il suo *opposto*), ovvero per ogni $a \in \mathbb{Z}$ esiste un $b \in \mathbb{Z}$ tale che

$$a + b = b + a = 0$$

(infatti, $b = -a$).

(5) *il prodotto è associativo*, ovvero per ogni $a, b, c \in \mathbb{Z}$ si ha

$$(ab)c = a(bc)$$

(6) *il prodotto è commutativo*, ovvero per ogni $a, b \in \mathbb{Z}$ si ha

$$ab = ba$$

(7) *esiste un elemento neutro per il prodotto* (il numero 1), tale che per ogni $a \in \mathbb{Z}$ si ha

$$a1 = 1a = a$$

(8) *vale la proprietà distributiva del prodotto rispetto alla somma*, ovvero per ogni $a, b, c \in \mathbb{Z}$ si ha

$$a(b + c) = ab + ac, \quad (a + b)c = ac + bc$$

Si noti che negli interi non vale la proprietà analoga alla (4) per il prodotto: infatti, non è vero che ogni elemento $a \in \mathbb{Z}$ ha inverso rispetto al prodotto, che dovrebbe essere un $b \in \mathbb{Z}$ tale che $ab = ba = 1$ (ad esempio, non esiste nessun numero intero b tale che $2b = 1$).
Tale proprietà è invece verificata per tutti gli $a \neq 0$ nell'insieme $\mathbb{Q}$ dei numeri razionali.

Osservazione 3.1. Un insieme dotato di due operazioni che verificano tutte le proprietà (1)-(8) di sopra si dice *anello commutativo* (semplicemente *anello* se valgono tutte tranne la (6)).
Un anello commutativo in cui ogni elemento diverso da 0 ha un inverso rispetto al prodotto si dice un *campo*: l'insieme dei razionali $\mathbb{Q}$ è quindi un esempio di campo (ad esempio, 2 ha come inverso moltiplicativo $\frac{1}{2}$ in $\mathbb{Q}$); un altro esempio è dato dall'insieme $\mathbb{R}$ dei numeri reali e da quello $\mathbb{C}$ dei numeri complessi (che vedremo nel capitolo successivo).
Nella matematica e nelle sue applicazioni si incontrano vari esempi di anelli diversi dall'insieme degli interi $\mathbb{Z}$ (ne vedremo altri nei prossimi paragrafi e capitoli): lavorare con questa definizione astratta permette di dimostrare una volta per tutte in generale risultati e formule che poi saranno validi in ogni anello, senza doverli dimostrare caso per caso (esattamente come già visto con la nozione di gruppo: ad esempio il Lemma 2.34, dimostrato una volta per tutte, sarà valido per ogni gruppo che incontreremo).

Osservazione 3.2. Il fatto che l'anello $\mathbb{Z}$ si estenda al campo $\mathbb{Q}$, cioè in altre parole il fatto che vedendo gli interi come particolari razionali essi ammettano

(in $\mathbb{Q}$) un inverso moltiplicativo di ogni numero diverso da zero ha come conseguenza che in $\mathbb{Z}$ vale la cosiddetta *legge di annullamento del prodotto*:

$$ab = 0 \Rightarrow a = 0 \vee b = 0$$

cioè se un prodotto è nullo allora almeno uno dei due fattori deve essere nullo[1].
Infatti, sia $ab = 0$ e supponiamo per assurdo che a e b siano entrambi diversi da zero: essendo gli interi particolari razionali, possiamo pensare l'uguaglianza $ab = 0$ anche come un'uguaglianza tra razionali e quindi, moltiplicando entrambi i membri per $\frac{1}{a}$ (che nei razionali esiste), troviamo da una parte $\frac{1}{a}(ab) = \frac{1}{a}0 = 0$, dall'altra

$$\frac{1}{a}(ab) = (\frac{1}{a}a)b = 1b = b$$

ovvero $b = 0$, che contraddice l'ipotesi che a e b fossero entrambi nulli.
Sottolineiamo che tale proprietà non va data per scontata, in quanto vedremo nei paragrafi e capitoli successivi contesti in cui essa non vale.

Ora, una cosa fondamentale che si può fare nell'anello degli interi $\mathbb{Z}$ e che useremo continuamente in questo capitolo è *eseguire la divisione con resto*. Più precisamente, vale il seguente

Teorema 3.3. Siano $a, b \in \mathbb{Z}$, con $b \neq 0$. Allora esiste un'unica coppia (q, r) di interi tali che valgono le due seguenti condizioni:

$$a = bq + r \tag{3.1}$$

$$0 \leq r < |b| \tag{3.2}$$

(dove $|b|$ indica il cosiddetto *valore assoluto di* b, uguale a b se b è positivo o nullo e a $-b$ se b è negativo: in altre parole se necessario si cambia il segno di b per renderlo non negativo). L'intero q si dice *quoziente*, l'intero r *resto* della divisione.

Esempio 3.4. Ad esempio, siano $a = 7$, $b = 2$. Si ha

$$7 = 2 \cdot 3 + 1$$

[1]Proprietà usata spesso quando si risolvono le equazioni, ad esempio per risolvere $x^2 - x = 0$ osservo che $x^2 - x = x(x-1)$, e quindi la mia equazione si riscrive $x(x-1) = 0$ che per la legge di annullamento del prodotto è verificata se $x = 0$ o $x - 1 = 0$ (ovvero $x = 1$).

quindi quoziente e resto della divisione sono rispettivamente $q = 3$ e $r = 1$. Si noti che ad esempio si ha anche

$$7 = 2 \cdot 2 + 3$$

oppure

$$7 = 2 \cdot 4 + (-1)$$

ma queste due uguaglianze, pur essendo del tipo (3.1), non verificano la (3.2), in quanto nella prima il resto 3 non è minore di $b = 2$, nella seconda il resto -1, pur essendo minore di $|b| = 2$ non è maggiore o uguale a zero.
Questo esempio mostra anche che affinché quoziente e resto siano unici e ben determinati una volta dati a e b, la condizione (3.2) è necessaria.

Osservazione 3.5. Come enunciato nel Teorema 3.3, a e b possono essere anche negativi: ad esempio se $a = -7$ e $b = 2$, si ha

$$-7 = 2 \cdot (-4) + 1$$

oppure se $a = 7$ e $b = -2$ si ha

$$7 = (-2) \cdot (-3) + 1$$

Un'applicazione importante dell'algoritmo della divisione è la possibilità di *scrivere un qualunque numero intero positivo in una qualsiasi base* $b \geq 2$.
Prima di dare la definizione precisa, osserviamo che quando scriviamo un qualunque numero in notazione decimale, ad esempio $n = 1375$, quello che le cifre $1, 3, 7, 5$ fanno è dirci (a partire da destra a sinistra) quante sono le unità, quante le decine, quante le centinaia etc., ovvero

$$1375 = 1 \cdot 1000 + 3 \cdot 100 + 7 \cdot 10 + 5$$

ovvero, usando la notazione usuale per le potenze

$$1375 = 1 \cdot 10^3 + 3 \cdot 10^2 + 7 \cdot 10^1 + 5$$

Quindi qualunque numero non negativo può essere scritto come somma di potenze decrescenti di 10 (l'ultima, a moltiplicare 5, sarebbe $10^0 = 1$ e non la scriviamo) precedute da un coefficiente che può essere un qualunque numero tra 0 e 9 (ovvero un qualunque numero minore di 10 e maggiore o uguale a zero). Vedremo ora che, usando l'algoritmo della divisione, la stessa cosa può essere fatta con un qualunque altro numero naturale maggiore o uguale di 2 al posto di 10:

Teorema 3.6. Sia $b \geq 2$ un intero fissato. Allora ogni intero positivo n può essere scritto in modo unico nella forma seguente

$$n = r_N b^N + r^{N-1} b^{N-1} + \cdots + r_1 b + r_0 \tag{3.3}$$

con tutti gli r_i, $i = 0, 1, \ldots, N$ tali che $0 \leq r_i < b$ (e $r_N \neq 0$). Si scrive allora $n = (r_N\ r_{N-1}\ \cdots\ r_1\ r_0)_b$ e si dice che $(r_N\ r_{N-1}\ \cdots\ r_1\ r_0)_b$ è la *scrittura di n in base b*.

Osservazione 3.7. La condizione che il primo coefficiente r_N sia diverso da zero serve a evitare addendi inutili nella (3.3) e garantire l'unicità della scrittura in base b. Ad esempio, 1375 è anche uguale a $0 \cdot 10^4 + 1 \cdot 10^3 + 3 \cdot 10^2 + 7 \cdot 10^1 + 5$, che vuol dire che potremmo anche scriverlo in notazione decimale come 01375, ma chiaramente il primo 0 si può anche omettere.

Come abbiamo detto, per trovare la scrittura di un numero n dato in base b, si usa l'algoritmo della divisione. Più precisamente, il procedimento è il seguente: si divide prima n per b ottenendo

$$n = q_0 b + r_0 \tag{3.4}$$

Se q_0 è minore di b, abbiamo finito: la (3.4) sarebbe già una combinazione di potenze di b (in questo caso compaiono solo $b = b^1$ e, sottointesa, $b^0 = 1$ che motiplica r_0) con coefficienti q_0 e r_0 minori di b (r_0 lo è sicuramente per definizione di resto), come prevede il Teorema 3.6.
Se invece $q_0 \geq b$, allora dividiamo q_0 per b:

$$q_0 = q_1 b + r_1 \tag{3.5}$$

e sostituiamo la (3.5) nella (3.4), ottenendo

$$n = (q_1 b + r_1) b + r_0 = q_1 b^2 + r_1 b + r_0 \tag{3.6}$$

Di nuovo, se q_1 è minore di b, abbiamo finito: la (3.6) sarebbe proprio una combinazione di potenze di b (in questo caso compaiono b^3, b^2, $b^1 = b$ e il termine di grado zero) con coefficienti q_1, r_1, r_0 minori di b.
Se invece $q_1 \geq b$, iteriamo il procedimento: dividiamo q_1 per b:

$$q_1 = q_2 b + r_2 \tag{3.7}$$

e sostituiamo la (3.7) nella (3.6), ottenendo

$$n = (q_2 b + r_2) b^2 + r_1 b + r_0 = q_2 b^3 + r_2 b^2 + r_1 b + r_0 \tag{3.8}$$

e se q_3 è minore di b, abbiamo finito.

Dal momento che i quozienti successivi che otteniamo sono sempre più piccoli, prima o poi arriveremo sicuramente a un quoziente minore di b, e il procedimento si arresta sicuramente[2].

Esempio 3.8. Ad esempio, scriviamo il numero $n = 19$ in base $b = 2$ con il procedimento descritto sopra: iniziamo dividendo 19 per 2:

$$19 = 9 \cdot 2 + 1$$

Poichè il quoziente ottenuto $q_0 = 9$ è maggiore di 2, eseguiamo la seconda divisione:

$$9 = 4 \cdot 2 + 1$$

e sostituiamo

$$19 = (4 \cdot 2 + 1) \cdot 2 + 1 = 4 \cdot 2^2 + 1 \cdot 2 + 1 \tag{3.9}$$

Questa non à ancora l'espressione di 19 in base 2 in quanto l'ultimo quoziente ottenuto, $q_1 = 4$, non è ancora minore di 2; dividendo ulteriormente otteniamo

$$4 = 2 \cdot 2 + 0$$

e sostitutendo al posto di 4 nella (3.9) si ottiene

$$19 = (2 \cdot 2 + 0) \cdot 2^2 + 1 \cdot 2 + 1 = 2 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2 + 1 \tag{3.10}$$

Non abbiamo ancora finito in quanto l'ultimo quoziente ottenuto, $q_2 = 2$, non è minore di 2: eseguiamo allora un'ulteriore divisione per 2

$$2 = 1 \cdot 2 + 0$$

finalmente, l'ultimo quoziente ottenuto, $q_3 = 1$, è minore di 2: sostituendo nella (3.10) si ottiene allora

$$19 = (1 \cdot 2 + 0) \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2 + 1 = 1 \cdot 2^4 + 0 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2 + 1 \tag{3.11}$$

che è finalmente l'uguaglianza cercata, che esprime 19 come combinazione di potenze di 2 con coefficienti minori di 2 (cioè 0 o 1). Possiamo anche scrivere che

$$19 = (1\ 0\ 0\ 1\ 1)_2$$

[2]Si noti che se avessimo ammesso come base $b = 1$ questo non sarebbe vero: dividendo n per 1 si ottiene $n = 1 \cdot n + 0$ ovvero quoziente n, che rividiso per 1 dà nuovamente quoziente n, e il procedimento non si arresta mai.

ovvero, più semplicemente, che 19 in base 2 si scrive come

$$10011$$

Si noti che viceversa dato un numero in una certa base, è facile verificare di quale numero si tratti in notazione decimale: ad esempio, in base 2 il numero 1101 rappresenta, in base alla definizione data nella formula (3.3),

$$1 \cdot 2^3 + 1 \cdot 2^2 + 0 \cdot 2 + 1 = 13$$

La numerazione in base $b = 2$ (o *numerazione binaria*) è di particolare importanza nell'informatica in quanto, esprimendo qualunque numero come successione di 0 e 1, consente a un computer di registrarlo o esprimerlo come successione di stati spento/acceso. Storicamente, altri casi importanti sono $b = 16$ (numerazione *esadecimale*) o $b = 8$ (numerazione *ottale*); tuttavia, come afferma il Teorema 3.6, si può usare qualunque $b \geq 2$: ad esempio, se volessimo scrivere 23 in base 7, eseguiremmo la divisione

$$23 = 3 \cdot 7 + 2$$

e poiché il quoziente ottenuto $q_1 = 3$ è già minore della base $b = 7$, il procedimento si ferma qui: 23, in base 7, si scrive semplicemente come $(3\ 2)_7$.

3.2 Divisori e numeri primi

Se, dividendo un intero a per un intero b capita che il resto sia zero, ovvero che

$$a = b \cdot q$$

allora si dice che a *è un multiplo di* b, o equivalentemente che b *è un divisore di* a (o anche che b *divide* a): in formule, si scrive $b|a$.
Ad esempio, sia $a = 20$. È facile verificare[3] che i divisori di 20 sono

$$\pm 1, \pm 20, \pm 2, \pm 5, \pm 4, \pm 10$$

(si noti che se b è un divisore di a, ovvero se $a = bq$ per qualche q, allora anche $-b$ è divisore di a in quanto vale chiaramente anche $a = (-b)(-q)$).
Si osservi che tra i divisori di $a \in \mathbb{Z}$ ci sono sicuramente ± 1 (in quanto $a = 1 \cdot a$ e $a = (-1) \cdot (-a)$) e $\pm a$ (in quanto $a = a \cdot 1$ e $a = (-a) \cdot (-1)$). I restanti divisori si chiamano *divisori propri*.
I numeri che non hanno divisori propri sono di fondamentale importanza nella matematica e in tutte le sue applicazioni:

[3] Vedremo dopo un metodo generale.

Definizione 3.9. Un numero intero $p \neq \pm 1$ si dice *primo* se i suoi unici divisori sono ± 1 e $\pm p$.

Come vedremo, i numeri primi sono di grandissima importanza per gli scopi di questo corso, quindi è importante conoscerne le proprietà.
Ad esempio, una prima proprietà, che caratterizza i numeri primi, è la seguente:

$$p|ab \Rightarrow p|a \vee p|b$$

ovvero se un numero primo divide un prodotto, esso divide necessariamente uno dei fattori.
Per convincersi che tale proprietà non vale per numeri in generale, si noti ad esempio che 10 divide il prodotto $4 \cdot 15 = 60$ ma non è vero né che 10 divide 4 né che 10 divide 15. Il motivo è che essendo $10 = 2 \cdot 5$, esso divide $4 \cdot 15$ in quanto $2|4$ e $5|15$: per un numero primo p che divida un prodotto ab una cosa del genere non è possibile in quanto non avendo divisori propri non può accadere che si scomponga in un prodotto in cui un fattore divide a e uno divida b.
La proprietà più importante dei numeri primi è però sicuramente quella espressa nel seguente risultato, detto *teorema fondamentale dell'aritmetica*:

Teorema 3.10. Dato un qualunque intero $a \neq 0, \pm 1$, esistono numeri primi $p_1, p_2, \ldots, p_s$ tali che

$$a = p_1 p_2 \cdots p_s$$

e tale decomposizione è unica a meno del segno e dell'ordine dei fattori.

Il teorema afferma in un certo senso che i primi sono i "mattoni" di cui si compongono tutti i numeri interi.
Ad esempio, una decomposizione in primi di $a = 20$ è

$$20 = 2 \cdot 2 \cdot 5$$

Come afferma il teorema, qualunque altre decomposizione di 20 in primi differisce da questa solo per il segno e per l'ordine dei fattori, ad esempio

$$20 = 2 \cdot 5 \cdot 2$$

oppure

$$20 = (-2) \cdot (-5) \cdot 2.$$

La condizione che i fattori della decomposizione siano primi è indispensabile per avere l'unicità a meno dell'ordine e del segno: ad esempio $20 = 5 \cdot 4$ e

$20 = 2{\cdot}10$ sono due decomposizioni in fattori (non primi) di 20 che differiscono per i fattori e non solo per l'ordine o il segno.

Osservazione 3.11. Spesso il teorema fondamentale dell'aritmetica si trova nei testi formulato come segue: dato un intero $a > 1$, esiste un'unica decomposizione

$$a = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_s^{\alpha_s}$$

in primi positivi $p_1, p_2, \ldots, p_s$ tali che $p_1 > p_2 > \cdots > p_s$, con gli esponenti $\alpha_1, \alpha_2, \ldots, \alpha_s > 0$. Rispetto al Teorema 3.10, ci stiamo limitando agli interi positivi (e quindi la condizione $a > 1$ che serve a escludere $a = 0$ e $a = 1$, come nel Teorema 3.10), il che consente di limitarsi ai primi positivi e evitare l'unicità meno del segno, mentre l'unicità a meno dell'ordine non viene più menzionata in quanto aggiungendo la condizione che i primi distinti che compaiono nella decomposizione siano disposti dal più grande al più piccolo stiamo in pratica fissando noi tale ordine.

In tale forma, la decomposizione di un intero in fattori primi ci permette di trovare velocemente tutti i suoi divisori: essi saranno tutti quelli del tipo $b = \pm p_1^{\beta_1} p_2^{\beta_2} \cdots p_s^{\beta_s}$, con gli esponenti $0 \leq \beta_1 \leq \alpha_1$, $0 \leq \beta_2 \leq \alpha_2$, $\ldots$, $0 \leq \beta_s \leq \alpha_s$.
Ad esempio, per $20 = 5^1 \cdot 2^2$, i divisori positivi sono tutti e soli i numeri del tipo $5^{\beta_1} 2^{\beta_2}$ con $0 \leq \beta_1 \leq 1$ e $0 \leq \beta_2 \leq 2$, ovvero

$$5^0 2^0 = 1, \quad 5^0 2^1 = 2, \quad 5^0 2^2 = 4, \quad 5^1 2^0 = 5, \quad 5^1 2^1 = 10, \quad 5^1 2^2 = 20$$

Dati due numeri interi a e b, questi possono chiaramente avere dei divisori in comune: ad esempio, è facile verificare che $a = 30$ e $b = 45$ hanno come divisori comuni

$$\pm 1, \pm 3, \pm 5, \pm 15$$

Notiamo che tutti questi divisori comuni dividono 15 (o, in altre parole, 15 è multiplo di tutti i divisori comuni): 15 si chiama allora *massimo comune divisore* di 30 e 45:

Definizione 3.12. Siano a e b due interi non entrambi nulli. Un intero positivo d si dice *massimo comune divisore di a e b* se sono soddisfatte entrambe le seguenti condizioni:

(1) d è un divisore comune di a e b (in simboli, $d|a \wedge d|b$)

(2) ogni altro divisore comune di a e b divide d (in simboli, $c|a \wedge c|b \Rightarrow c|d$)

Osservazione 3.13. La condizione che a e b siano non entrambi nulli serve a garantire l'esistenza del massimo comune divisore: infatti, 0 ha come divisori tutti i numeri interi, in quanto $0 = n \cdot 0$ per ogni $n \in \mathbb{Z}$, quindi se $a = 0$ e $b = 0$ ogni intero sarebbe un divisore comune di a e b e non ce ne sarebbe uno massimo nel senso della Definizione 3.12[4].
Invece, la condizione che b sia positivo serve a garantire l'unicità: per esempio, nell'esempio di sopra con $a = 30$ e $b = 45$ anche -15 è un divisore comune che è multiplo di ogni altro divisore comune.

Il massimo comune divisore di due interi a e b può essere calcolato semplicemente scrivendo prima tutti i divisori di a, poi tutti quelli di b (usando il metodo descritto nell'Osservazione 3.11) e poi guardando quale tra tutti i divisori comuni è multiplo degli altri.
Tuttavia, tale metodo non è molto efficace in quanto richiede la conoscenza della scomposizione in fattori primi di a e b, che non sempre è facile da trovare. Vedremo quindi un altro metodo, basato sull'algoritmo della divisione, che come vedremo dopo ci dà anche altre informazioni.
Il metodo funziona come segue: si inizia dividendo a per b:

$$a = bq_1 + r_1 \tag{3.12}$$

Si divide poi b per il resto r_1

$$b = r_1 q_2 + r_2 \tag{3.13}$$

ottenendo un secondo resto r_2: a questo punto si divide il primo resto per il secondo resto

$$r_1 = r_2 q_3 + r_3 \tag{3.14}$$

il secondo resto per il terzo resto

$$r_2 = r_3 q_4 + r_4 \tag{3.15}$$

e così via, fino a che non si ottiene resto zero:

$$\vdots$$

$$r_n = r_{n+1} q_{n+2} + r_{n+2} \tag{3.16}$$

$$r_{n+1} = r_{n+2} q_{n+3} \tag{3.17}$$

Affermiamo che l'ultimo resto r_{n+2} non nullo di queste divisioni successive è esattamente il massimo comune divisore di a e b.
Prima di mostrare perché, vediamo alcuni esempi:

[4]Invece uno solo tra a e b può essere nullo, ad esempio se $a = 0$ e $b > 0$, i divisori comuni di a e b sono tutti i divisori di b, e il massimo sarebbe b stesso.

Esempio 3.14. Siano $a = 45$ e $b = 30$, come sopra. Dividendo a per b otteniamo

$$45 = 30 \cdot 1 + 15$$

e dividendo $b = 30$ per il resto $r_1 = 15$ otteniamo

$$30 = 15 \cdot 2$$

Quindi l'ultimo resto prima di ottenere resto zero è 15, che risulta essere come avevamo già detto sopra il massimo comune divisore.
Vediamo ora $a = 42$ e $b = 30$: dividendo a per b otteniamo

$$42 = 30 \cdot 1 + 12$$

Dividendo $b = 30$ per il resto $r_1 = 12$ si ha

$$30 = 12 \cdot 2 + 6$$

e dividendo $r_1 = 12$ per il secondo resto $r_2 = 6$ si ottiene

$$12 = 6 \cdot 2$$

cioè resto zero: l'ultimo resto non nullo, 6, è quindi il massimo comune divisore.
Infine, siano $a = 120$ e $b = 23$: procedendo come sopra, si ha

$$120 = 23 \cdot 5 + 5$$

$$23 = 5 \cdot 4 + 3$$

$$5 = 3 \cdot 1 + 2$$

$$3 = 2 \cdot 1 + 1$$

$$2 = 1 \cdot 2$$

L'ultimo resto non nullo è 1, che risulta essere quindi il massimo comune divisore.

L'ultimo esempio di sopra è un caso della seguente, importante

Definizione 3.15. Due interi a e b si dicono *primi tra loro* o *coprimi* se il loro massimo comune divisore è 1.

In altre parole, due interi sono coprimi se non hanno divisori comuni a parte il caso banale 1.

Vediamo ora perchè il procedimento per divisioni successive descritto e illustrato fornisce effettivamente il massimo comune divisore. Dobbiamo dimostrare che l'ultimo resto non nullo r_{n+2} che compare nella (3.16) non solo è un divisore comune di a e b, ma che ogni altro divisore comune di a e b divide r_{n+2}.

Per dimostrare la prima affermazione, osserviamo che l'ultima divisione svolta, la (3.17) (quella con resto nullo) ci dice che r_{n+2} divide r_{n+1} o, equivalentemente, r_{n+1} è un multiplo di r_{n+2}. Sostituendola nella (3.16), si trova

$$r_n = r_{n+1}q_{n+2} + r_{n+2} = r_{n+2}q_{n+3}q_{n+2} + r_{n+2} = r_{n+2}(q_{n+3}q_{n+2} + 1) \quad (3.18)$$

che ci dice che r_{n+2} divide anche r_n o, equivalentemente, r_n è un multiplo di r_{n+2}: per brevità, riscriviamo la (3.18) come $r_n = r_{n+2}k$.

Ora, la divisione immediatamente precedente alla (3.16) sarà

$$r_{n-1} = r_n q_{n+1} + r_{n+1} \quad (3.19)$$

Ma sapendo che sia r_n che r_{n+1} sono multipli di r_{n+2}, ovvero $r_{n+1} = r_{n+2}q_{n+3}$ e $r_n = r_{n+2}k$, sostituendo nella (3.19) otteniamo

$$r_{n-1} = r_n q_{n+1} + r_{n+1} = r_{n+2}kq_{n+1} + r_{n+2}q_{n+3} = r_{n+2}(kq_{n+1} + q_{n+3}) \quad (3.20)$$

il che dimostra che r_{n+2} divide anche r_{n-1}: in altre parole, stiamo mostrando, risalendo lungo tutte le divisioni effettuate dalla prima all'ultima, che r_{n+2} divide i vari resti r_{n+1}, r_n, r_{n-1} etc.: quando saremo arrivati alla (3.15) otterremo che r_{n+2} divide r_2, la (3.14) ci dirà che r_{n+2} divide r_1, e infine la (3.13) ci dirà che r_{n+2} divide b e la (3.12) ci dirà che r_{n+2} divide a. Quindi r_{n+2} è un divisore comune di a e b.

Per dimostrare che si tratta effettivamente del massimo comune divisore, basta mostrare che r_{n+2} è multiplo di qualunque divisore comune di a e b. Sia allora c un tale divisore comune: quindi possiamo scrivere $a = ca'$ e $b = cb'$. Sostituendo queste due uguaglianze nella (3.12), otteniamo

$$ca' = cb'q_1 + r_1$$

ovvero, portando il primo addendo a primo membro e mettendo in evidenza c,

$$c(a' - b'q_1) = r_1 \quad (3.21)$$

Questa uguaglianza ci dice che il primo resto r_1 è anche lui un multiplo di c: per brevità, riscriviamo la (3.21) come $r_1 = cr_1'$. Sostituendo questa uguaglianza e la $b = cb'$ nella (3.13), si ottiene

$$cb' = cr_1'q_2 + r_2$$

ovvero, portando il primo addendo a primo membro e mettendo in evidenza c,

$$c(b' - r_1'q_2) = r_2 \tag{3.22}$$

Questa uguaglianza ci dice che anche il secondo resto r_2 è un multiplo di c: continuando in questo modo per sostituzioni successive nelle divisioni, dalla prima all'ultima, vediamo che tutti i resti sono un multiplo di c, fino ad arrivare a r_{n+2}, il che dimostra che r_{n+2} è proprio il massimo comune divisore, come volevamo.

Vediamo ora una importante proprietà del massimo comune divisore, che ci sarà molto utile nel paragrafo successivo:

Proposizione 3.16. Sia d il massimo comune divisore di due interi a e b. Allora, esistono due interi $x, y \in \mathbb{Z}$ tali che

$$d = ax + by \tag{3.23}$$

Non dimostriamo la proposizione, ma ci limitiamo a mostrare come trovare gli interi x e y in alcuni esempi concreti.
Consideriamo il terzo caso considerato nell'Esempio 3.14 in cui, tramite la successione di divisioni

$$120 = 23 \cdot 5 + 5$$

$$23 = 5 \cdot 4 + 3$$

$$5 = 3 \cdot 1 + 2$$

$$3 = 2 \cdot 1 + 1$$

$$2 = 1 \cdot 2$$

abbiamo mostrato che il massimo comune divisore di 120 e 23 è $d = 1$ (l'ultimo resto non nullo delle divisioni). Allo scopo di trovare i due interi x e y tali che $120x + 23y = 1$, la cui esistenza è prevista nella Proposizione 3.16, iniziamo con l'osservare che se nella penultima divisione $3 = 2 \cdot 1 + 1$ portiamo il primo addendo a primo membro, ovvero

$$1 = 3 - 2 \cdot 1 = 1 \cdot 3 + (-1) \cdot 2 \tag{3.24}$$

otteniamo che siamo riusciti a riscrivere il massimo comune divisore come combinazione degli ultimi due resti 2 e 3 ottenuti prima del massimo comune divisore stesso. Ora, combinando la (3.24) con la divisione precedente $5 = 3 \cdot 1 + 2$ possiamo scrivere 1 come combinazione dei penultimi due resti 5 e 3: più precisamente, portando a primo membro nella $5 = 3 \cdot 1 + 2$ otteniamo $2 = 5 - 3 \cdot 1$ che sostituito al posto di 2 nella (3.24) ci dà

$$1 = 1 \cdot 3 + (-1) \cdot (5 - 3 \cdot 1) = 3 - 5 + 3 = 3 \cdot 2 + (-1) \cdot 5. \tag{3.25}$$

Continuando con questa idea, usando la divisione $23 = 5 \cdot 4 + 3$, riscritta come $3 = 23 - 5 \cdot 4$, e sostituendo al 3 che compare nella (3.25) si ottiene

$$\begin{aligned} 1 = 3 \cdot 2 + (-1) \cdot 5 = (23 - 5 \cdot 4) \cdot 2 + (-1) \cdot 5 = \\ = 23 \cdot 2 - 5 \cdot 4 \cdot 2 + (-1) \cdot 5 = 23 \cdot 2 + (-9) \cdot 5 \end{aligned} \tag{3.26}$$

cioè siamo arrivati a scrivere il massimo comune divisore come combinazione di 23 e 5.
A questo punto, possiamo usare la prima divisione $120 = 23 \cdot 5 + 5$, riscritta $5 = 120 - 23 \cdot 5$, per esprimere nella combinazione (3.26) il 5 in funzione di 120 e 23, ovvero, sostituendo,

$$\begin{aligned} 1 = 23 \cdot 2 + (-9) \cdot 5 = 23 \cdot 2 + (-9) \cdot (120 - 23 \cdot 5) = \\ = 23 \cdot 2 + (-9) \cdot 120 + (-9) \cdot (-5) \cdot 23 = \\ 47 \cdot 23 + (-9) \cdot 120 \end{aligned} \tag{3.27}$$

e siamo riusciti finalmente a esprimere 1 nella forma $23x + 120y$, con $x = 47$ e $y = -9$.

Vediamo un ulteriore esempio: sia $a = 42$ e $b = 30$, per i quali abbiamo già mostrato che il massimo comune divisore è 6 nell'Esempio 3.14 mediante la successione di divisioni

$$42 = 30 \cdot 1 + 12$$

$$30 = 12 \cdot 2 + 6$$

$$12 = 6 \cdot 2$$

Come sopra, usiamo la penultima divisione $30 = 12 \cdot 2 + 6$ per esprimere il massimo comune divisore $6 = 30 + (-2) \cdot 12$ come combinazione di 30 e 12; in tale uguaglianza sostituiamo $12 = 42 - 30$, ricavata dalla prima divisione, ottenendo

$$6 = 30 + (-2) \cdot (42 - 30)$$

. Sviluppando i conti, otteniamo

$$6 = 30 + (-2) \cdot 42 + 2 \cdot 30 = 3 \cdot 30 + (-2) \cdot 42$$

cioè siamo riusciti, come volevamo, a esprimere 6 nella forma $30x + 42y$, con $x = 3$ e $y = -2$.

Osservazione 3.17. Se a e b sono interi uno multiplo dell'altro, diciamo a è multiplo di b, allora il procedimento descritto sopra non si applica: ad esempio se $a = 140$ e $b = 14$, la prima divisione $140 = 14 \cdot 10 + 0$ dà già resto zero, e quindi non possiamo usare il fatto che il massimo comune divisore è l'ultimo resto non nullo delle divisioni successive. Tuttavia, in tal caso il massimo comune divisore è semplicemente uguale a b, che in tale situazione è chiaramente divisore di a e anche di b, ed è chiaramente il massimo; anche per ottenere la forma $d = ax+by$ non è necessario applicare nessun procedimento di divisioni, in quanto essendo $d = b$ basta scrivere $d = a0 + b1$.

Concludiamo questo paragrafo con la seguente domanda: dato un intero positivo N, quanti sono gli interi positivi $k < N$ coprimi con N (ovvero quelli che hanno come unico divisore in comune con N l'intero 1)?
Il numero di tali interi si denota con $\phi(N)$, e definisce quindi una funzione ϕ detta *funzione di Eulero*, che è di fondamentale importanza nelle applicazioni che vedremo nell'ultimo paragrafo di questo capitolo.
Vediamo allora come calcolare $\phi(N)$ per ogni N: l'idea è sfruttare il fatto che N, in base al teorema fondamentale dell'aritmetica, si può scrivere come $N = P_1^{\alpha_1} P_2^{\alpha_2} \cdots P_s^{\alpha_s}$ con $P_1, P_2, \ldots, P_s$ primi.
Innanzitutto, se N è un numero primo P, si ha

$$\phi(P) = P - 1 \tag{3.28}$$

Infatti, dal momento che P è primo, per definizione esso non ha altri divisori (positivi) oltre 1 e P stesso. Quindi, un numero k, per avere divisori diversi da 1 in comune con P dovrebbe essere un multiplo di P, ma questo chiaramente non può succedere se $k < P$: quindi tutti gli interi $k < P$, ovvero $k = 1, 2, \ldots, P - 1$ sono coprimi con P, da cui la (3.28).
Consideriamo ora il caso in cui $N = P^\alpha$ è potenza di un primo: si ha

$$\phi(P^\alpha) = P^\alpha - P^{\alpha-1} \tag{3.29}$$

Infatti, $\phi(P^\alpha)$ è dato dai numeri positivi minori di P^α che sono coprimi con P^α, che possiamo pensare come tutti i numeri da 1 a P^α, che sono proprio P^α, meno quelli che hanno divisori in comune (oltre a 1) con P^α: se dimostriamo

allora che gli interi minori o uguali a P^α che hanno divisori in comune con P^α sono $P^{\alpha-1}$, avremo finito.
Dal momento che P^α è potenza di P, un intero k può avere un divisore in comune con P^α solo se contiene P tra i suoi fattori, ovvero solo se k è un multiplo di P. Tali multipli sono chiaramente

$$P,\ 2P,\ 3P, \ldots, P^{\alpha-1}P$$

(ci siamo fermati a $P^{\alpha-1}P = P^\alpha$ perché dobbiamo considerare solo interi minori o uguali a P^α), e quindi il numero di questi interi non coprimi con P^α è $P^{\alpha-1}$, come volevamo.

Infine, per calcolare $\phi(N)$ per ogni N useremo il seguente risultato, che non dimostriamo.

Lemma 3.18. Siano N_1 e N_2 due interi primi tra loro: allora

$$\phi(N_1 N_2) = \phi(N_1)\phi(N_2) \tag{3.30}$$

Possiamo combinare questo lemma con la formula (3.29) per calcolare finalmente la funzione di Eulero per ogni N.
Supponiamo che sia, in base al teorema fondamentale dell'aritmetica, $N = P_1^{\alpha_1} P_2^{\alpha_2} \cdots P_s^{\alpha_s}$.
Chiaramente i fattori $P_1^{\alpha_1}$, $P_2^{\alpha_2}, \ldots, P_s^{\alpha_s}$ sono tutti primi tra loro in quanto potenze di primi diversi: quindi, in base al Lemma 3.18, ho

$$\phi(N) = \phi(P_1^{\alpha_1} P_2^{\alpha_2} \cdots P_s^{\alpha_s}) = \phi(P_1^{\alpha_1})\phi(P_2^{\alpha_2}) \cdots \phi(P_s^{\alpha_s}) = \tag{3.31}$$

(in base alla (3.29))

$$= (P_1^{\alpha_1} - P_1^{\alpha_1 - 1})(P_2^{\alpha_2} - P_2^{\alpha_2 - 1}) \cdots (P_s^{\alpha_s} - P_s^{\alpha_s - 1}) \tag{3.32}$$

Se nelle parentesi tonde che compaiono nella (3.32) mettiamo in evidenza rispettivamente $P_1^{\alpha_1}, P_2^{\alpha_2}, \ldots, P_s^{\alpha_s}$, possiamo riscriverla come

$$\phi(N) = \left[P_1^{\alpha_1}\left(1 - \frac{1}{P_1}\right)\right]\left[P_2^{\alpha_2}\left(1 - \frac{1}{P_2}\right)\right] \cdots \left[P_s^{\alpha_s}\left(1 - \frac{1}{P_s}\right)\right] \tag{3.33}$$

ovvero, permutando i fattori,

$$\phi(N) = P_1^{\alpha_1} P_2^{\alpha_2} \cdots P_s^{\alpha_s}\left(1 - \frac{1}{P_1}\right)\left(1 - \frac{1}{P_2}\right) \cdots \left(1 - \frac{1}{P_s}\right) \tag{3.34}$$

che, ricordando che $P_1^{\alpha_1} P_2^{\alpha_2} \cdots P_s^{\alpha_s}$ è proprio N, significa

$$\phi(N) = N\left(1 - \frac{1}{P_1}\right)\left(1 - \frac{1}{P_2}\right)\cdots\left(1 - \frac{1}{P_s}\right) \tag{3.35}$$

che è la formula che si trova solitamente nei libri per ϕ.
Ad esempio, se $N = 100$, abbiamo $N = 5^2 \cdot 2^2$, quindi

$$\phi(100) = 100\left(1 - \frac{1}{5}\right)\left(1 - \frac{1}{2}\right) = 100 \cdot \frac{4}{5} \cdot \frac{1}{2} = 40.$$

Osserviamo che il calcolo di ϕ mediante la formula (3.35) richiede di conoscere la decomposizione di N in fattori primi, che non è sempre facile da trovare.

Osservazione 3.19. Lo studio dei numeri primi, e in particolare della loro distribuzione nell'insieme dei numeri interi, costituisce un capitolo fondamentale nella matematica. Intanto, sappiamo che i numeri primi sono infiniti: se per assurdo fossero in numero finito, e fossero $p_1, p_2, \ldots, p_k$, il numero $N = p_1 p_2 \cdots p_k + 1$ sarebbe un numero non divisibile per nessun primo (il resto della divisione di N per uno qualunque dei p_i sarebbe sempre 1), ma questo per il teorema fondamentale dell'aritmetica non è possibile a meno che non sia anche lui primo: quindi, anche N è primo, contraddicendo l'ipotesi che i primi fossero solo $p_1, p_2, \ldots, p_k$.
Come è distribuito tale insieme infinito all'interno degli interi? quanto è raro trovare primi quando si considerano numeri sempre più grandi?
Da una parte, si congettura che esistano infinite coppie di numeri primi a distanza 2 l'uno dall'altro (ad esempio 3 e 5, oppure 17 e 19), detti *numeri primi gemelli*, dall'altra esistono sequenze lunghe quanto vogliamo di numeri che non contengono nessun primo: dato n, la sequenza

$$n! + 2, n! + 3, \ldots, n! + n$$

è una sequenza di $n - 1$ numeri consecutivi (quindi possiamo renderla lunga quanto vogliamo scegliendo n grande) in cui nessuno è primo: infatti, $n!+2$ è divisibile per 2 in quanto sia $n!$ che 2 lo sono (ricordiamo che $n!$ è il prodotto di tutti i numeri naturali da 1 a n), $n! + 3$ è divisibile per 3 in quanto sia $n!$ che 3 lo sono, e così via.
Un importante teorema afferma che se $\pi(n)$ indica il numero dei primi minori di n, allora la quantità $\frac{\pi(n)}{n}$ (che descrive la percentuale di primi minori di n rispetto al totale dei numeri da 1 a n) si avvicina sempre di più a $\frac{1}{\log n}$ (ovvero il rapporto tra $\frac{\pi(n)}{n}$ e $\frac{1}{\log n}$ si avvicina sempre di più a 1 al crescere di n).