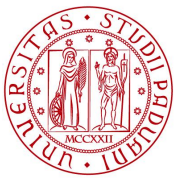

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Università degli Studi di Padova

DIPARTIMENTO DI MATEMATICA "TULLIO LEVI-CIVITA"

Corso di Laurea Triennale in Matematica

Trattazione di alcune classi
particolari di equazioni diofantee

Relatore:
Prof. Matteo Longo

Laureando: Mattia Oberto
Matricola: 2017596

Anno Accademico 2022/2023

data 12/12/2023

Indice

Introduzione

Lo scopo di questa tesi di laurea è quello di raccogliere alcuni strumenti per permettere di risolvere le equazioni diofantee, ovvero equazioni polinomiali a coefficienti interi di cui cercheremo le soluzioni intere o razionali, di diversi gradi. In particolare nella prima parte tratteremo delle classi di equazioni diofantee più semplici, ovvero quelle lineari e quadratiche, e nella seconda parte approfondiremo lo studio di alcune equazioni diofantee cubiche. Vedremo infatti che le diofantee cubiche sono la prima classe di equazioni diofantee di cui non riusciremo sempre a determinare le soluzioni con i nostri strumenti, e saranno quindi più complesse da studiare delle precedenti. Per permetterci di apprendere più informazioni riguardo a queste ultime equazioni dimostreremo il Teorema di Mordell, che ci darà informazioni determinanti per lo studio delle soluzioni razionali dell'equazione, e vedremo come sfruttare il processo dimostrativo di questo teorema per ottenere, in alcuni casi, le soluzioni che cerchiamo.

Parte I

Casi noti di equazioni diofantee

Capitolo 1

Equazioni diofantee lineari

In questo capitolo e nel prossimo ci occuperemo di trovare le soluzioni di alcune equazioni diofantee più semplici, ovvero di equazioni polinomiali in più variabili a coefficienti interi e di grado 1 e 2. In particolare ne cercheremo le soluzioni intere.

Definizione 1.1 Un'equazione scritta nella forma $f(x_1, \ldots, x_n) = 0$, in cui $f \in \mathbb{Z}[\, x_1, \ldots, x_n]$, per $n \geq 2$ è detta *equazione diofantea (algebrica).*

Definizione 1.2 Un'equazione diofantea scritta nella forma

$$a_1x_1 + \ldots + a_nx_n = c$$

con $a_i \in \mathbb{Z}$, per $i = 1, \ldots, n$ è detta *equazione diofantea lineare.*

Richiamiamo ora un fatto noto e la definizione di congruenza modulo n, per $n \in \mathbb{Z}$, che utilizzeremo spesso per le equazioni diofantee.

Proposizione 1.1 *$\mathbb{Z}$ è un dominio euclideo, la cui funzione grado $\delta : \mathbb{Z}\backslash\{0\} \to \mathbb{N}$ è definita da $\delta(a) = |a|,\ \forall a \in \mathbb{Z}\backslash\{0\}$.*

Definizione 1.3 Siano $a, b, n \in \mathbb{Z}$. Allora a, b si dicono *congrui modulo n* se $a - b \in n\mathbb{Z}$ e si indica con

$$a \equiv b \pmod{n}.$$

La nozione di congruenza modulo n e divisibilità sono strettamente legate: se consideriamo ad esempio $a, n \in \mathbb{Z}$, con $n \neq 0$, allora $\exists\ q, r \in \mathbb{Z}$ tali che $a = nq + r$ (Proposizione (1.1)) con $|r| < |n|$. Ora $a - r = nq \ \in n\mathbb{Z} \ \Rightarrow \ a \equiv r \pmod{n}$, quindi ogni intero a è sempre congruo al resto della divisione euclidea per n, modulo $n \in \mathbb{Z}\backslash\{0\}$. Segue che

$$n|a \iff a \equiv 0 \pmod{n}.$$

Teorema 1.2
Siano $a, b, c \in \mathbb{Z}$, con $a, b \neq 0$. Consideriamo l'equazione diofantea lineare

$$ax + by = c \tag{1.1}$$

allora:

1. *l'equazione* (1.1) *ha soluzione in* $\mathbb{Z} \iff d := MCD(a,b)$ *divide* c.

2. *Se* (x_0, y_0) *è una soluzione particolare dell'equazione* (1.1)*, allora ogni soluzione intera è nella forma*
$$(x,y) = \left(x_0 + \frac{b}{d}t,\ y_0 - \frac{a}{d}t\right)$$
al variare di $t \in \mathbb{Z}$.

Dimostrazione.
1. $\Rightarrow$) Sia (x_0, y_0) una soluzione dell'equazione (1.1), ovvero soddisfa l'uguaglianza
$$ax_0 + by_0 = c.$$
Osserviamo che $d|a$ e $d|b$, quindi in particolare dividerà una qualunque loro combinazione lineare, per cui vale $d|ax_0 + by_0 = c$.
$\Leftarrow$) Per ipotesi sappiamo che $d|a, b, c$, quindi esistono $a_1, b_1, c_1 \in \mathbb{Z}$ t.c.
$$\begin{cases} a = da_1 \\ b = db_1 \\ c = dc_1 \end{cases}$$
Sostituiamo le relazioni appena trovate nell'equazione (1.1) e dividiamola per d. Ora l'equazione sarà nella forma
$$a_1x + b_1y = c_1$$
con $MCD(a_1, b_1) = 1$. Essendo a_1 e b_1 primi tra loro, b_1 ammette inverso moltiplicativo modulo a_1. Detto k tale inverso moltiplicativo, riduciamo l'intera equazione modulo a_1 e moltiplichiamola per k. Otteniamo
$$kc_1 = \ kb_1y = \ y$$
Scegliendo quindi $y_0 = kc_1$ otteniamo
$$x_0 = \frac{c_1 - b_1x_0}{a_1} = \ c_1\frac{1 - b_1k}{a_1}$$
ed essendo $b_1k \equiv 1 \pmod{a_1}$ abbiamo che $a_1|1 - b_1k \ \Rightarrow \ x_0 \in \mathbb{Z}$.
Dunque $a_1x_0 + b_1y_0 = c_1$, per $x_0, y_0 \in \mathbb{Z}$, ovvero (x_0, y_0) è soluzione intera dell'equazione (1.1).
2. Data (x_0, y_0) soluzione dell'equazione (1.1), allora
$$\left(x_0 + \frac{b}{d}t, y_0 - \frac{a}{d}t\right) \text{ al variare di } t \in \mathbb{Z}$$
è ancora soluzione. Infatti
$$\begin{aligned} a\left(x_0 + \frac{b}{d}t\right) + b\left(y_0 - \frac{a}{d}t\right) &= ax_0 + \frac{abt}{d} + by_0 - \frac{bat}{d} \\ &= ax_0 + by_0 = c. \end{aligned}$$

Viceversa siano $h, k \in \mathbb{Z}$ t.c. $(x_0 + h, y_0 + k)$ sia soluzione dell'equazione (1.1). Allora

$$\begin{aligned} c &= a(x_0 + h) + b(y_0 + k) \\ &= ax_0 + by_0 + ah + bk \\ &= c + ah + bk \end{aligned}$$

Da cui ricaviamo $k = -\frac{ah}{b}$, e quindi $\frac{ah}{b} \in \mathbb{Z} \Rightarrow b|ah$.

Ora dati a_1, b_1 definiti come nella dimostrazione del punto *1.*, $db_1|da_1h \Rightarrow b_1|a_1h$. Ma $MCD(a_1, b_1) = 1$, quindi $b_1|h$. Ciò significa che esiste un intero t tale per cui $h = tb_1 = \frac{b}{d}t$, da cui

$$k = -\frac{ah}{b} = -\frac{a}{d}t.$$

Quindi ogni soluzione dell'equazione (1.1) è nella forma

$$\begin{cases} x = x_0 + \frac{b}{d}t \\ y = y_0 - \frac{a}{d}t \end{cases} .$$

□

Esempio 1.1 Troviamo le coppie (x, y) di soluzioni intere tali che

$$19x + 31y = 13.$$

Soluzione : riducendo l'equazione modulo 19 otteniamo

$$12y \equiv 13 \pmod{19}$$

Potremmo procedere trovando l'inverso moltiplicativo di 12 modulo 19 procedendo per tentativi, ma sarebbe un processo laborioso che possiamo invece semplificare: riscriviamo invece l'informazione ottenuta dalla congruenza modulo 19: $\exists\ k \in \mathbb{Z}\ t.c.$

$$12y = 13 + 19k.$$

In questo modo otteniamo la nuova equazione diofantea lineare

$$12y - 19k = 13.$$

Avendo scelto di guardare l'equazione modulo 19, siamo riusciti a ridurre la grandezza del coefficiente delle y e ad ottenere così coefficienti più piccoli che richiedono calcoli più semplici. Per cui ora guardiamo iteriamo il procedimento riducendo l'equazione modulo 12, ottenendo

$$-7k \equiv 1 \pmod{12}.$$

Il problema si è quindi ridotto a trovare l'inverso moltiplicativo di 7 modulo 12, che richiede un numero massimo di tentativi più basso rispetto al precedente. Iterando il procedimento è possibile quindi ridurre ulteriormente i coefficienti con cui si lavora. Nel nostro caso ci fermiamo a questo passaggio notando che $49 \equiv 1 \pmod{12}$ e quindi che

possiamo scegliere $k = 5$ essendo $-7k \equiv 1 \equiv 49 \pmod{12}$ vale che $k \equiv -7 \equiv 5 \pmod{12}$. Sostituendo nelle equazioni precedenti otteniamo

$$y = \frac{13 + 19 \cdot 5}{12} = 9$$

e quindi

$$x = \frac{13 - 31 \cdot 9}{19} = -14$$

Da cui la soluzione particolare $(-14, 9)$ e tutte e sole le soluzioni $(-14 + 31j, 9 - 19j)$ al variare di $j \in \mathbb{Z}$.

Osservazione 1.1. Il problema stesso di trovare l'inverso moltiplicativo di un intero dato un modulo fissato è riducibile alla risoluzione di un'equazione diofantea lineare in due variabili e quindi a trovare inversi moltiplicativi sempre più semplici. Posso infatti trovare l'inverso moltiplicativo k di $a \in \mathbb{Z}$ modulo $b \in \mathbb{Z}$ cercando le soluzioni intere (k, j) dell'equazione

$$ak - bj = 1.$$

La risoluzione di tali equazioni diofantee solitamente avviene attraverso l'algoritmo di Euclide, che sfrutta la divisione euclidea e che ci permette di ottenere delle soluzioni.

Esempio 1.2 (Algoritmo di Euclide) Consideriamo l'equazione diofantea lineare vista nell'Esempio (1.1):

$$31x + 19y = 13.$$

Cerchiamone le soluzioni intere (x, y) con l'algoritmo di Euclide.

Definiamo una successione $(r_n)_{n \geq -1}$ nel seguente modo: I primi termini della successione sono

$$r_{-1} := 31 = a, \quad r_0 := 19 = b.$$

Per i termini successivi, finché $r_i \neq 0$ possiamo avvalerci della divisione euclidea per riscrivere il termine r_{i-1} e ottenere il termine r_{i+1}. In particolare, fissati $r_{i-1}, r_i \in \mathbb{Z}$, $r_i \neq 0$, $i \geq 0$, esistono degli interi $q_i, r_{i+1} \in \mathbb{Z}$ tali che:

$$r_{i-1} = r_i q_i + r_{i+1}, \text{ per } 0 \leq r_{i+1} < r_i$$

e ci fermiamo quando otteniamo per la prima volta $r_i = 0$. Chiamiamo $N \in \mathbb{N}$ l'indice per il quale $r_{N+1} = 0$. Osserviamo che

$$\begin{aligned} MCD(r_{i-1}, r_i) &= MCD(r_{i-1} - q_i r_i, r_i) \\ &= MCD(r_i q_i + r_{i+1} - r_i q_i, r_i) \\ &= MCD(r_{i+1}, r_i) \end{aligned}$$

per $i = 0, \ldots, N-1$. Di conseguenza

$$\begin{aligned} MCD(19, 31) &= MCD(r_{-1}, r_0) = \ldots = MCD(r_N, r_{N+1}) \\ &= MCD(r_N, 0) = r_N = 1. \end{aligned}$$

Nel nostro caso otteniamo

$$\begin{aligned} 31 = 19 \cdot 1 + 12 &\Rightarrow q_0 = 1,\ r_1 = 12 \\ 19 = 12 \cdot 1 + 7 &\Rightarrow q_1 = 1,\ r_2 = 7 \\ 12 = 7 \cdot 1 + 5 &\Rightarrow q_2 = 1,\ r_3 = 5 \\ 7 = 5 \cdot 1 + 2 &\Rightarrow q_3 = 1,\ r_4 = 2 \\ 5 = 2 \cdot 2 + 1 &\Rightarrow q_4 = 2,\ r_5 = 1 \\ 2 = 1 \cdot 2 + 0 &\Rightarrow q_5 = 2,\ r_6 = 0. \end{aligned}$$

Ripercorrendo a ritroso il procedimento è possibile trovare due interi h, k tali che $31h + 19k = 1$ e, di conseguenza, $(13h, 13k)$ sarà una soluzione intera dell'equazione, essendo che $31 \cdot 13h + 19 \cdot 13k = 13(31h + 19k) = 13$. Procediamo definendo

$$\begin{aligned} x_{-1} = y_0 = 1, \\ x_0 = y_{-1} = 0, \end{aligned}$$

e poi ricorsivamente

$$\begin{aligned} x_j = x_{j-2} - q_{j-1}x_{j-1} \\ y_j = y_{j-2} - q_{j-1}y_{j-1}. \end{aligned}$$

Dimostriamo ora che vale la relazione $ax_j + by_j = r_j$ per induzione sull'indice j.
$\mathbf{j = -1}$: $ax_{-1} + by_{-1} = a = r_{-1}$.
$\mathbf{j = 0}$: $ax_0 + by_0 = b = r_0$.
$\mathbf{n \geq 1}$: vale

$$ax_{j-2} + by_{j-2} - q_{j-1}(ax_{j-1} + by_{j-1}) \overset{*}{=} r_{j-2} - q_{j-1}r_{j-1} = r_j$$

in cui l'uguaglianza $*$ vale per ipotesi induttiva. In particolare quindi $ax_n + by_n = r_N = 1$. Esplicitamente otteniamo:

$$\begin{aligned} x_1 = 1 - 1 \cdot 0 = 1, \quad & y_1 = 0 - 1 \cdot 1 = -1 \\ x_2 = -1, \quad & y_2 = 2 \\ x_3 = 2, \quad & y_3 = -3 \\ x_4 = -3, \quad & y_4 = 5 \\ x_5 = 8, \quad & y_5 = -13 \\ x_6 = -19, \quad & y_6 = 31 \end{aligned}$$

otteniamo così la soluzione $13(x_5, y_5) = (104, -169)$ perché, come ci aspettiamo, vale $31 \cdot 8 + 19 \cdot (-13) = 248 - 247 = 1$. Per il Teorema (1.2), possiamo trovare le altre soluzioni come $(104 - 19t, -169 + 31t)$ al variare di $t \in \mathbb{Z}$.

Enunciamo e dimostriamo ora il Teorema cinese del resto nel caso di un sistema di congruenze su $\mathbb{Z}$.

Teorema 1.3 *(Teorema cinese del resto)*
Siano dati $n_1, \dots, n_k \in \mathbb{Z}$ *a due a due coprimi con* $n_i \geq 2 \ \forall i \in \{1, \dots, k\}$ *e siano* $a_1, \dots, a_k \in \mathbb{Z}$.
Allora il sistema di congruenze:

$$\begin{cases} x \equiv a_1 \pmod{n_1} \\ \vdots \\ x \equiv a_k \pmod{n_k} \end{cases}$$

ammette soluzione. Inoltre tale soluzione è unica modulo $\prod_{i=1}^{k} n_i$.

Dimostrazione.
Sia $P := \prod_{i=1}^{k} n_i$. Dimostriamo che esiste una biezione naturale π tra gli insiemi

$$\mathbb{Z}/P\mathbb{Z} \longleftrightarrow \{(a_1, \dots, a_k) : a_i \in \mathbb{Z}/n_i\mathbb{Z} \ \forall i \in \{1, \dots, k\}\}$$

che associa all'elemento $x \in \mathbb{Z}/P\mathbb{Z}$ la k-upla $(a_1, \dots, a_k)$ t.c. $\begin{cases} x \equiv a_1 \pmod{n_1} \\ \vdots \\ x \equiv a_k \pmod{n_k} \end{cases}$.

Notiamo che, essendo

$$|\mathbb{Z}/P\mathbb{Z}| = |\{(a_1, \dots, a_k) : a_i \in \mathbb{Z}/n_i\mathbb{Z} \ \forall i \in \{1, \dots, k\}\}| =: P$$

se l'applicazione è iniettiva allora è anche suriettiva.

Ne dimostriamo l'iniettività: siano $x_1, x_2 \in \mathbb{Z}/P\mathbb{Z}$ t.c. $\pi(x_1) = \pi(x_2) = (a_1, \dots, a_k)$. Allora

$$x_1 \equiv x_2 \pmod{n_i}, \ \forall i \in \{1, \dots, k\}$$

o, equivalentemente, $n_i | x_1 - x_2 \ \forall i \in \{1, \dots, k\}$. Sfruttiamo il fatto che essendo $n_1, \dots, n_k$ a due a due coprimi, vale

$$P = \prod_{i=1}^{k} n_i = mcm(n_1, \dots, n_k)$$

e la condizione appena trovata per affermare che $P = mcm(n_1, \dots, n_k) | x_1 - x_2$, cioè $x_1 \equiv x_2 \pmod{P}$. Ma $x_1, x_2 \in \mathbb{Z}/P\mathbb{Z}$, quindi la congruenza modulo P dei due numeri è un'uguaglianza, ovvero $x_1 = x_2$. Ciò dimostra l'iniettività e quindi la suriettività della mappa.

Essendo π biettiva, è invertibile con inversa π^{-1}. Inoltre dalla suriettività di π segue che per ogni sistema di congruenze

$$\begin{cases} x \equiv a_1 \pmod{n_1} \\ \vdots \\ x \equiv a_k \pmod{n_k} \end{cases}$$

esiste $x \in \mathbb{Z}/P\mathbb{Z}$ dato da $x = \pi^{-1}(a_1, \ldots, a_k)$, che sarà soluzione del sistema.

Concludiamo osservando che per l'iniettività tale x è unico modulo P.

□

Esempio 1.3 Troviamo $x \in \mathbb{Z}$ tale che

$$\begin{cases} x \equiv 1 \pmod 6 \\ x \equiv 7 \pmod{28} \end{cases}$$

I moduli non sono tra loro coprimi, quindi sarà necessario riscrivere le due equivalenze modulo potenza di un primo:

$$\begin{cases} x \equiv 1 \pmod 2 \\ x \equiv 1 \pmod 3 \\ x \equiv 3 \pmod 4 \\ x \equiv 0 \pmod 7 \end{cases}$$

Notiamo ora che $x \equiv 3 \pmod 4 \Rightarrow x \equiv 1 \pmod 2$ e quindi l'informazione fornita dalla prima congruenza è, in questo caso, superflua. Ora il sistema si è ridotto quindi a sole tre congruenze con moduli a due a due coprimi, per cui sappiamo che esiste un'unica soluzione x modulo $3 \cdot 4 \cdot 7 = 84$ del sistema:

$$\begin{cases} x \equiv 1 \pmod 3 \\ x \equiv 3 \pmod 4 \\ x \equiv 0 \pmod 7 \end{cases}$$

In questo caso la soluzione $x = 7$ soddisfa tutte le condizioni del sistema, per cui tutte e sole le soluzioni intere del sistema saranno della forma $x = 7 + 84k$ al variare di $k \in \mathbb{Z}$.

Osservazione 1.2. La condizione di coprimità a due a due dei moduli è necessaria per affermare in generale l'esistenza di soluzioni: nell'Esempio (1.3) avevamo due informazioni diverse che riguardavano x modulo potenza di 2. Il fatto che le due informazioni fossero consistenti non vale in generale. Ad esempio il sistema:

$$\begin{cases} x \equiv 1 \pmod 6 \\ x \equiv 8 \pmod{28} \end{cases}$$

ci dice che $x \equiv 1 \pmod 2$ e che $x \equiv 0 \pmod 4$. Quindi x è dispari per la prima informazione ma divisibile per 4 per la seconda, il che è assurdo.

Sussiste una stretta relazione tra equazioni diofantee lineari e il Teorema cinese del resto. Consideriamo ad esempio un sistema di 2 congruenze

$$\begin{cases} x \equiv a \pmod m \\ x \equiv b \pmod n \end{cases}$$

con $MCD(m,n)=1$. Cerchiamo allora $k,j\in\mathbb{Z}$ t.c.

$$\begin{cases} x-a=km \\ x-b=jn \end{cases}, \text{ ovvero } \begin{cases} x=a+km \\ x=b+jn \end{cases}$$

Da cui ricaviamo

$$a+km=b+jn \;\Rightarrow\; mk-nj=b-a.$$

Quest'ultima è un'equazione diofantea lineare nelle variabili k e j, ed essendo $MCD(m,-n)$ $=MCD(m,n)=1$, ammette sempre soluzione. Abbiamo quindi mostrato che dal Teorema (1.2) otteniamo l'esistenza di soluzioni di un sistema di 2 congruenze. In particolare dall'esistenza della soluzione per un sistema di due congruenze è dimostrabile induttivamente l'esistenza di soluzioni per un sistema di k congruenze con moduli a due a due coprimi. Infatti:
Passo base : $\mathbf{k=2}$. Segue dall'osservazione precedente.
Passo induttivo : L'ipotesi induttiva in questo caso è l'esistenza di soluzioni per sistemi di k con moduli a due a due coprimi. Cerchiamo ora soluzioni per il sistema

$$\begin{cases} x\equiv a_1 \pmod{n_1} \\ \vdots \\ x\equiv a_k \pmod{n_k} \\ x\equiv a_{k+1} \pmod{n_{k+1}} \end{cases}.$$

Consideriamo le ultime due condizioni del sistema. Equivalgono a un sistema di due congruenze con moduli a due a due coprimi che ammette una soluzione b_k modulo $n_k\cdot n_{k+1}$ per il Passo base della dimostrazione. Segue che il sistema iniziale è equivalente al sistema:

$$\begin{cases} x\equiv a_1 \pmod{n_1} \\ \vdots \\ x\equiv a_{k-1} \pmod{n_{k-1}} \\ x\equiv b_k \pmod{n_k\cdot n_{k+1}} \end{cases}$$

ma $MCD(n_i,n_k\cdot n_{k+1})|MCD(n_i,n_k)\cdot MCD(n_i,n_{k+1})=1$, quindi i moduli del nuovo sistema di k congruenze sono ancora a due a due coprimi, e quindi ammette soluzione per ipotesi induttiva.

Similmente a come appena fatto per il Teorema Cinese del Resto possiamo estendere la teoria sulle equazioni diofantee lineari per più di due variabili in maniera induttiva, a partire dal Teorema (1.2).

Teorema 1.4
Siano $a_1,\dots,a_n,c\in\mathbb{Z}$. Allora l'equazione

$$a_1x_1+\dots+a_nx_n=c \tag{1.2}$$

ammette soluzione $(x_1,\dots,x_n)\iff MCD(a_1,\dots,a_n)|c$.

Dimostrazione.
$\Rightarrow$) Sia $d = MCD(a_1, \ldots, a_n)$. Allora $d|a_i$ per $i = 1, \ldots, n$ e quindi $d|a_1x_1 + \ldots + a_nx_n = c$.
$\Leftarrow$) **Passo base** : $\mathbf{n = 2}$. Segue dal Teorema (1.2).
Passo induttivo : per ipotesi induttiva sappiamo che dati $b_1, \ldots, b_n \in \mathbb{Z}$, allora $MCD(b_1, \ldots, b_n)|k \Rightarrow$ l'equazione $b_1x_1 + \ldots + b_nx_n = k$ ammette soluzioni intere.

Consideriamo l'equazione

$$a_1x_1 + \ldots + a_{n+1}x_{n+1} = c \tag{1.3}$$

con $MCD(a_1, \ldots, a_{n+1})|c$, e l'equazione

$$a_nx_n + a_{n+1}x_{n+1} = d_n \tag{1.4}$$

nella quale $d_n := MCD(a_n, a_{n+1})$, che ammette soluzione per il Teorema (1.2). Sia (y_n, y_{n+1}) una soluzione dell'equazione (1.4). Introduciamo una variabile $z_n \in \mathbb{Z}$ tale che

$$\begin{cases} x_n = y_nz_n \\ x_{n+1} = y_{n+1}z_n \end{cases}.$$

In questo modo l'equazione (1.3) diventa

$$a_1x_1 + \ldots + a_{n-1}x_{n-1} + (a_ny_n + a_{n+1}y_{n+1})z_n = a_1x_1 + \ldots + a_{n-1}x_{n-1} + d_nz_n = c \tag{1.5}$$

che è un'equazione diofantea lineare nelle variabili $(x_1, \ldots, x_{n-1}, z_n)$.

Dimostriamo adesso che $d := MCD(a_1, \ldots, a_n, a_{n+1}) = MCD(a_1, \ldots, a_{n-1}, d_n)$:

$$\begin{aligned} d|a_n,\ d|a_{n+1} &\Rightarrow d \mid MCD(a_n, a_{n+1}) = d_n \\ &\Rightarrow d \mid MCD(a_1, \ldots, a_{n-1}, d_n). \end{aligned}$$

Ora

$$\begin{cases} MCD(a_1, \ldots, a_{n-1}, d_n) \mid MCD(a_1, \ldots, a_{n-1}) \mid a_1, \ldots, a_{n-1} \\ MCD(a_1, \ldots, a_{n-1}, d_n) \mid d_n \mid a_n, a_{n+1} \end{cases}$$

Dunque $MCD(a_1, \ldots, a_{n-1}, d_n) \mid a_1, \ldots, a_{n+1}$. Ma allora vale

$$MCD(a_1, \ldots, a_{n-1}, d_n) \mid MCD(a_1, \ldots, a_{n+1}) = d,$$

quindi $MCD(a_1, \ldots, a_{n-1}, d_n) = d \mid c$.

Possiamo quindi concludere per ipotesi induttiva che l'equazione (1.5) ammette soluzioni intere $(x_1, \ldots, x_{n-1}, z_n)$ e, di conseguenza, l'equazione (1.3) ammette soluzioni intere $(x_1, \ldots, x_{n-1}, y_nz_n, y_{n+1}z_n)$. □

Esempio 1.4 Trovare tutte le terne di interi (x, y, z) che soddisfano l'equazione:

$$10x + 14y + 35z = 23. \tag{1.6}$$

Le ipotesi del Teorema (1.4) sono soddisfatte, infatti $MCD(10, 14, 35) = 1 \mid 23$. Troviamo allora le soluzioni dell'equazione:

$MCD(14, 35) = 7$, quindi risolviamo l'equazione:

$$14a + 35b = 7, \quad a, b \in \mathbb{Z}$$

o analogamente

$$2a + 5b = 1, \quad a, b \in \mathbb{Z} \tag{1.7}$$

le cui soluzioni sono $(3 - 5k, -1 + 2k)$ al variare di $k \in \mathbb{Z}$.

Consideriamo l'equazione

$$10x + 7h = 23, \quad h \in \mathbb{Z}. \tag{1.8}$$

Tutte e sole le sue soluzioni $(3 - 7j, -1 + 10j)$ al variare di $j \in \mathbb{Z}$. Ora moltiplichiamo per h l'equazione (1.6):

$$14ah + 35bh = 7h$$

in cui scegliamo i valori di h che soddisfano l'equazione (1.8), cioè quelli nella forma $h = 10j - 1$ al variare di $j \in \mathbb{Z}$. Le coppie $(y, z) = (ah, bh)$ che fanno parte di terne di soluzioni sono quindi nella forma

$$((-1 + 10j)(3 - 5k),\ (-1 + 10j)(-1 + 2k)), \quad k, j \in \mathbb{Z}.$$

Infine quindi le terne di soluzioni sono tutte e sole le terne della forma

$$(x, y, z) = (3 - 7j,\ (-1 + 10j)(3 - 5k),\ (-1 + 10j)(-1 + 2k)), \quad k, j \in \mathbb{Z}.$$

Capitolo 2

Equazioni diofantee quadratiche

Adesso che abbiamo terminato la trattazione di equazioni diofantee lineari, passiamo alla trattazione di equazioni diofantee quadratiche, per cui è noto un procedimento per la determinazione di soluzioni in alcuni casi.

Definizione 2.1 Un'equazione diofantea scritta nella forma

$$\sum_{i=0}^{n} a_{ii}x_i^2 + \sum_{\substack{i,j=0 \\ i\neq j}}^{n} a_{ij}x_i x_j + \sum_{i=0}^{n} b_i x_i + c \tag{2.1}$$

si dice *equazione diofantea quadratica*.

2.1 Terne ed n-uple pitagoriche

Trattiamo prima di tutto un caso noto di equazione diofantea quadratica, ovvero le terne pitagoriche. Cerchiamo quindi le terne di interi (x, y, z) soluzioni dell'equazione

$$x^2 + y^2 = z^2. \tag{2.2}$$

Se (x, y, z) è soluzione, allora anche (kx, ky, kz) è soluzione $\forall\ k \in \mathbb{Z}$. Inoltre, se (x, y, z) è soluzione, allora $MCD(x, y) = MCD(x, z) = MCD(y, z) = MCD(x, y, z)$. Infatti se supponessimo per assurdo ad esempio che $d = MCD(x, y) \nmid MCD(x, z)$, allora $d \nmid z$, ma $d|x,\ d|y\ \Rightarrow d^2|x^2,\ d^2|y^2\ \Rightarrow\ d^2\ |x^2 + y^2 = z^2\ \Rightarrow\ d \mid z$.
Per simmetria di x, y e z nell'equazione (a meno di un fattore -1) possiamo concludere che $MCD(x, y) = MCD(x, y, z)$.
Diciamo allora *primitive* le soluzioni (x, y, z) t.c. $MCD(x, y) = 1$.

Teorema 2.1
Ogni soluzione primitiva (x, y, z) dell'equazione

$$x^2 + y^2 = z^2$$

con y pari è nella forma

$$\begin{cases} x = m^2 - n^2 \\ y = 2mn \\ z = m^2 + n^2 \end{cases}$$

con $m, n \in \mathbb{Z},\ m > n,\ m + n$ dispari e inoltre $MCD(m, n) = 1$.

Dimostrazione.
x e y non possono essere entrambi dispari, altrimenti $z^2 = x^2 + y^2 \equiv 2 \pmod 4$, ma allora $2|z^2 \Rightarrow 2|z \Rightarrow 4|z^2 \Rightarrow z^2 \equiv 0 \pmod 4$, che è una contraddizione. Quindi esattamente uno tra x e y è pari. Senza perdere di generalità, assumiamo sia $y = 2a,\ a \in \mathbb{Z}$. Ora x e z sono necessariamente entrambi dispari, per cui esistono $b, c \in \mathbb{Z}$ tali che $z + x = 2b,\ z - x = 2c$.

Osserviamo che

$$\begin{aligned} d := MCD(b, c) &= MCD\left(\frac{z+x}{2}, \frac{z-x}{2}\right) \\ &= MCD\left(z, \frac{z-x}{2}\right) \\ &= MCD\left(x, \frac{z-x}{2}\right), \end{aligned}$$

quindi $d|z$, $d|x$, ma x e z sono coprimi, quindi $d = 1$. Inoltre

$$4a^2 = y^2 = z^2 - x^2 = 4\left(\frac{z-x}{2}\right)\left(\frac{z+x}{2}\right) = 4bc$$

cioò $a^2 = bc$, ma essendo b e c coprimi devono essere necessariamente entrambi quadrati perfetti. per cui esistono $m, n \in \mathbb{Z}$ tali che $b = m^2,\ c = n^2$. Otteniamo:

$$\begin{cases} x = b - c = m^2 - n^2 \\ y = \sqrt{4bc} = 2mn \\ z = b + c = m^2 + n^2 \end{cases}$$

Concludiamo osservando che

$$(m^2 - n^2)^2 + (2mn)^2 = m^4 + 2m^2n^2 + n^4 = (m^2 + n^2)^2$$

quindi ogni terna in questa forma è soluzione per l'equazione (2.2). □

Possiamo trovare famiglie di soluzioni simili a quelle delle terne pitagoriche, anche per quaterne (x, y, z, t) tali che $x^2 + y^2 + z^2 = t^2$, infatti è facile verificare che le quaterne della forma

$$\begin{cases} x = l^2 + m^2 - n^2 \\ y = 2lm \\ z = 2mn \\ t = l^2 + m^2 + n^2 \end{cases}$$

sono soluzioni per $l, m, n \in \mathbb{Z}$, ma non sono le sole soluzioni dell'equazione. Vediamo quindi come è possibile generalizzare il problema delle terne pitagoriche con n-uple tali che $\sum_{i=1}^{n-1} x_i^2 = x_n^2$, per $n \geq 3$ trovandone tutte le soluzioni possibili nel seguente Teorema.

Teorema 2.2
Tutte e sole le soluzioni intere dell'equazione

$$x_1^2 + \ldots + x_k^2 = x_{k+1}^2 \tag{2.3}$$

con $MCD(x_1, \ldots, x_k) = 1$ *sono nella forma*

$$\begin{cases} x_1 = m_1 m_k \\ x_2 = m_2 m_k \\ \vdots \\ x_{k-1} = m_{k-1} m_k \\ x_k = \frac{(m_1^2 + m_2^2 + \ldots + m_{k-1}^2 - \alpha^2 m_k^2)}{2\alpha} \\ x_{k+1} = \frac{(m_1^2 + m_2^2 + \ldots + m_{k-1}^2 + \alpha^2 m_k^2)}{2\alpha} \end{cases} \tag{2.4}$$

per $m_1, \ldots, m_k \in \mathbb{Z}$ *con* $MCD(x_1, \ldots, x_k) = 1$.

Dimostrazione.
Mostriamo innanzitutto che ogni $(k+1)-$upla in questa forma è soluzione:

$$\begin{aligned} x_{k+1}^2 - x_k^2 &= (x_{k+1} - x_k)(x_{k+1} + x_k) \\ &= \frac{(2\alpha^2 m_k^2) \cdot \left(2 \sum_{i=1}^{k-1} m_i^2\right)}{4\alpha^2} \\ &= m_k^2 (m_1^2 + m_2^2 + \ldots + m_{k-1}^2) \\ &= x_1^2 + \ldots + x_{k-1}^2. \end{aligned}$$

Dimostriamo ora ogni soluzione $(x_1, \ldots, x_{k+1})$ di interi per l'equazione (2.3) con $MCD(x_1, \ldots, x_k) = 1$ è nella forma del sistema (2.4).

Sia $m_k = MCD(x_1, \ldots, x_{k-1})$. Riscriviamo $x_i = m_i m_k$ per $i = 1, \ldots, k-1$, scegliendo m_k a meno di un fattore 2. Allora

$$\sum_{i=1}^{k-1} x_i^2 = m_k^2 \sum_{i=1}^{k-1} m_i^2 = x_{k+1}^2 - x_k^2 = (x_{k+1} - x_k)(x_{k+1} + x_k).$$

Notiamo che $MCD(x_k, m_k) = MCD(x_{k+1}, m_k) = 1$, altrimenti sarebbe contraddetta la coprimalità degli x_i. Calcoliamo adesso

$$\begin{aligned} MCD(x_{k+1} - x_k,\ x_{k+1} + x_k) &= MCD(x_{k+1} - x_k,\ 2x_k) \\ &= tMCD(x_{k+1} - x_k,\ x_k) \\ &= tMCD(x_{k+1},\ x_k) =: td \end{aligned}$$

in cui $t \in \{1, 2\}$. Osserviamo che $MCD(m_k, d) = 1$, per cui se m_k^2 divide $(x_{k+1} - x_k)(x_{k+1} + x_k)$, deve dividere esattamente uno solo dei due fattori. Analizziamo entrambi i casi contemporaneamente e vedremo che sarà indifferente da un passaggio in particolare in poi effettuare la distinzione nei due casi:

$$\begin{aligned} x_{k+1} \mp x_k &= \alpha m_k^2 \\ x_{k+1} \pm x_k = x_{k+1} \mp x_k \pm 2x_k &= \alpha m_k^2 \pm 2x_k \\ x_{k+1}^2 - x_k^2 &= \alpha^2 m_k^4 \pm 2\alpha m_k^2 x_k \\ x_{k+1}^2 = x_k^2 + m_k^2(\alpha^2 m_k^2 \pm 2\alpha x_k) &= x_k^2 + m_k^2(m_1^2 + \ldots + m_{k-1}^2) \end{aligned}$$

da cui si ricava

$$m_1^2 + \ldots + m_{k-1}^2 = \alpha^2 m_k^2 \pm 2\alpha x_k$$

e quindi esplicitando x_k

$$x_k = \pm \frac{m_1^2 + \ldots + m_{k-1}^2 - \alpha^2 m_k^2}{2\alpha}.$$

Possiamo allora scegliere il segno come preferiamo, essendo che i valori degli x_i possono essere scelti indifferentemente dal segno. Ora

$$\begin{aligned} \sum_{i=1}^{k} x_i^2 &= \frac{\sum_{i=1}^{k-1} m_i^4 + 2\sum_{\substack{i,j=1 \\ i \neq j}}^{k-1} m_i^2 m_j^2 + \alpha^4 m_k^2 - 2\alpha^2 m_k^2 \cdot \left(\sum_{i=1}^{k-1} m_i^2\right)}{4\alpha^2} + m_k^2 \left(\sum_{i=1}^{k-1} m_i^2\right) \\ &= \frac{\sum_{i=1}^{k-1} m_i^4 + 2\sum_{\substack{i,j=1 \\ i \neq j}}^{k-1} m_i^2 m_j^2 + 2\alpha^2 m_k^2 \cdot \left(\sum_{i=1}^{k-1} m_i^2\right) + \alpha^4 m_k^4}{4\alpha^2} \\ &= \left(\frac{\alpha^2 m_k^2 + \sum_{i=1}^{k-1} m_i^2}{2\alpha} \right)^2. \end{aligned}$$

Quindi

$$x_{k+1} = \frac{m_1^2 + \ldots + m_{k-1}^2 + \alpha^2 m_k^2}{2\alpha}.$$

□

Esempio 2.1 Cerchiamo tutte le coppie (x, y) di interi positivi che soddisfano l'equazione

$$x^2 + y^2 = 1997(x - y).$$

sapendo che 1997 è un numero primo.

Sfruttiamo le due seguenti relazioni:

$$(x + y)^2 + (x - y)^2 = 2(x^2 + y^2) = 2 \cdot 1997(x - y)$$

$$(x + y)^2 + (1997 - x + y)^2 = 1997^2.$$

Vediamo subito che $x + y \leq 1997$. Ci siamo così ricondotti alla ricerca di soluzioni di una terna pitagorica della forma $a^2 + b^2 = 1997^2$, con $a, b \in \{0, 1, \ldots, 1997\}$. Inoltre, essendo 1997 primo, $MCD(a, b) = 1$. Questo perché se $d := MCD(a, b)$, allora $d \mid a^2 + b^2 = 1997^2$ e quindi $d \in \{1, 1997\}$, ma $d = 1997$ se e solo se $a, b = 1997$ che non soddisfano l'uguaglianza. Cerchiamo allora soluzioni nella forma

$$\begin{cases} 1997 = m^2 + n^2 \\ a = 2mn \\ b = m^2 - n^2 \end{cases}$$

Osserviamo che $1997 \equiv 2 \pmod 5$ e i residui quadratici modulo 5 sono $0, 1, -1$, per cui le riduzioni di m^2 ed n^2 modulo 5 possono assumere solo questi tre valori e la loro somma deve essere 2. L'unica soluzione possibile la si ottiene per $m^2, n^2 \equiv 1 \pmod 5$, di conseguenza $m, n \equiv \pm 1 \pmod 5$. Analogamente ricaviamo $m, n \equiv \pm 1 \pmod 3$. Abbiamo troviamo 4 casi possibili modulo 15 per m ed n, ovvero $m, n \equiv 1, 4, 11, 14 \pmod{15}$. Concludiamo osservando che $m \geq n$, perché abbiamo assunto $b \geq 0$, e $\frac{1997}{2} \leq m^2 \leq 1997$, perché

$$1997 = m^2 + n^2 \leq 2m^2$$
$$1997 = m^2 + n^2 \geq m^2,$$

per cui le uniche scelte possibili per m sono $m = 34, 41, 44$.

L'unica soluzione accettabile che si ottiene verificando a mano i singoli casi è per $(m, n) = (34, 29)$, da cui $(a, b) = (1972, 315)$. Le uniche soluzioni quindi sono le coppie

$$(x, y) = (170, 145) \quad \text{e} \quad (x, y) = (1827, 145).$$

2.2 L'equazione di Pell

Definizione 2.2 Sia $D \in \mathbb{N}$ un non quadrato perfetto ed $m \neq 0$ un intero. Allora l'equazione:

$$x^2 - Dy^2 = m \tag{2.5}$$

a soluzioni intere nonnegative (x, y) è detta *equazione di Pell*.

Le prime equazioni di Pell vengono studiate già ai tempi dei greci, nel caso di $m = 1$, per trovare delle buone approssimazioni razionali di radici quadrate di interi. Teone di Smirna utilizzò il rapporto x_0/y_0 per approssimare $\sqrt{2}$, dove x_0 e y_0 sono le soluzioni dell'equazione $x^2 - 2y^2 = 1$. In generale

$$x^2 = dy^2 + 1 \;\Rightarrow\; \frac{x^2}{y^2} = d + \frac{1}{y^2}.$$

L'idea alla base del procedimento quindi consiste nel fatto che per y grandi, x_0/y_0 è una buona approssimazione per $\sqrt{d}$.

Osservazione 2.1. Il caso in cui D sia un quadrato perfetto si riconduce alla risoluzione di equazioni diofantee lineari, infatti se $D = k^2$ allora:

$$m = x^2 - k^2y^2 = (x - ky)(x + ky)$$

per cui le soluzioni sono tutte e sole le soluzioni accettabili date dai sistemi

$$\begin{cases} x - ky = j \\ x + ky = \frac{m}{j} \end{cases}$$

al variare di j tra i divisori di m.

Osservazione 2.2. Se (x, y) è una soluzione di $x^2 - dy^2 = 1$, allora $(x^2 + dy^2,\ 2xy)$ è anch'essa soluzione, infatti

$$1^2 = (x^2 - dy^2)^2 = (x^2 + dy^2)^2 - (2xy)^2 d.$$

Ora osserviamo che per $d, x, y \neq 0$ si ha $|x| < x^2 + dy^2$, quindi se esiste una soluzione (x, y) è possibile generare infinite soluzioni, che saranno tra loro distinte essendo che il primo termine della coppia cresce ad ogni passaggio.

Teorema 2.3 (Lagrange)
Se D è un intero positivo che non è un quadrato perfetto, l'equazione

$$u^2 - Dv^2 = 1 \tag{2.6}$$

ha infinite soluzioni negli interi nonnegativi, che saranno date dai termini della successione $(u_n, v_n)_{n\geq 0}$

$$u_{n+1} = u_1u_n + Dv_1v_n, \qquad v_{n+1} = v_1u_n + u_1v_n \tag{2.7}$$

dove $(u_0, v_0) = (1, 0)$ è la soluzione banale e (u_1, v_1) è la soluzione fondamentale, ovvero tale che $v_1 > 0$ sia minimo.

Dimostrazione.
Dimostriamo che l'equazione (2.6) ha una soluzione fondamentale.

Sia $c_1 > 1$ un intero. Mostriamo che $\exists\ t_1, w_1 \in \mathbb{N}_{>0}$ t.c.

$$|t_1 - w_1\sqrt{D}| < \frac{1}{c_1} \quad \text{e} \quad w_1 \leq c_1.$$

Sia $l_k = \lfloor k\sqrt{D}+1 \rfloor,\ k = 0, \dots, c_1$. Allora $0 < l_k - k\sqrt{D} \leq 1$. Essendo poi che $\sqrt{D} > 1$, $l_{k_1} = l_{k_2}$ se e solo se $k_1 = k_2$. Quindi cui ci sono c_1 intervalli della forma $\left(\frac{p-1}{c_1}, \frac{p}{c_1}\right)$ per $p = 1, \dots, c_1$, mentre ci sono $c_1 + 1$ numeri nella forma $l_k - k\sqrt{D}$ per $k = 0, \dots, c_1$. Ci saranno allora almeno 2 numeri in quella forma appartenenti a uno stesso intervallo, ovvero $\exists\ i, j, p \in \{0, 1, \dots, c_1\}$ con $i \neq j$ e $p \neq 0$ tali che

$$\frac{p-1}{c_1} < l_i - i\sqrt{D} \leq \frac{p}{c_i} \quad e \quad \frac{p-1}{c_1} < l_j - j\sqrt{D} \leq \frac{p}{c_i}$$

Dalle disuguaglianze allora ricaviamo che $|(l_j - l_i) - (j-i)\sqrt{D}| < \frac{1}{c_1}$. Ponendo $t_1 := |l_j - l_i|$ e $w_1 = |j - i|$ segue $|t_1 - w_1\sqrt{D}| < \frac{1}{c_1}$, e $w_1 \leq c_1$. Notiamo che $t_1 + w_1\sqrt{D} < 2w_1\sqrt{D} + 1$, quindi moltiplicando membro a membro le due disuguaglianze otteniamo

$$|t_1^2 - Dw_1^2| < 2\frac{w_1}{c_1}\sqrt{D} + \frac{1}{c_1} < 2\sqrt{D} + 1.$$

Scegliamo ora un intero positivo $c_2 > c_1$ tale che $|t_1 - w_1\sqrt{D}| > \frac{1}{c_2}$, in modo da ottenere t_2, w_2 interi positivi con $|t_2 - w_2\sqrt{D}| < \frac{1}{c_2}$ e $w_2 \leq c_2$.

Come prima otteniamo

$$|t_2^2 - Dw_2^2| < 2\sqrt{D} + 1 \quad e \quad |t_1 - t_2| + |w_1 - w_2| \neq 0.$$

Iterando questa procedura, otteniamo una successione di coppie distinte $(t_n, w_n)_{n\geq 1}$ che soddisfano la disuguaglianza $|t_n^2 - Dw_n^2| < 2\sqrt{D} + 1,\ \forall n \in \mathbb{N}_{\geq 1}$. Di conseguenza l'intervallo $(-2\sqrt{D} - 1, 2\sqrt{D} + 1)$ contiene al suo interno un intero $k \neq 0$ tale che esista una sottosuccessione della successione $(t_n, w_n)_{n\geq 1}$ che soddisfa l'equazione $t^2 - Dw^2 = k$. Questa sottosuccesione contiene almeno due coppie $(t_s, w_s),\ (t_r, w_r)$ per le quali

$$t_s \equiv t_r \pmod{|k|},$$

$$w_s \equiv w_r \pmod{|k|}$$

e $t_s w_r - t_r w_s \neq 0$, altrimenti $t_s = t_r$ e $w_s = w_r$, che contraddice $|t_s - t_r| + |w_s - w_r| \neq 0$. Sia $t_0 = t_s t_r - D w_s w_r$ e $w_0 = t_s w_r - t_r w_s$. Allora

$$\begin{aligned} t_0^2 - Dw_0^2 &= t_s^2 t_r^2 + D^2 w_s^2 w_r^2 - 2Dt_s t_r w_s w_r - Dt_s^2 w_r^2 - Dt_r^2 w_s^2 + 2Dt_s t_r w_s w_r \\ &= t_s^2(t_r^2 - Dw_r^2) - Dw_s^2(t_r^2 - Dw_r^2) \\ &= (t_s^2 - Dw_s^2)(t_r^2 - Dw_r^2) \\ &= k^2. \end{aligned}$$

Inoltre $t_0 = t_s t_r - D w_s w_r \equiv t_s^2 - Dw_0^2 \equiv 0 \pmod{|k|}$, e vale anche $w_0 \equiv 0 \pmod{|k|}$ di conseguenza. La coppia (t, w), ove $t_0 = t|k|$ e $w_0 = w|k|$ è una soluzione non banale dell'equazione (2.6).

Mostriamo ora che le coppie (u_n, v_n) definite come nella ricorsione (2.7) soddisfano l'equazione (2.6) per induzione su n. Abbiamo appena concluso la dimostrazione del passo

base per la coppia (u_1, v_1). Supponiamo quindi che (u_n, v_n) sia soluzione e dimostriamo che lo è anche (u_{n+1}, v_{n+1}).

$$\begin{aligned} u_{n+1}^2 - Dv_{n+1}^2 &= (u_1u_n + Dv_1v_n)^2 - D(v_1u_n + u_1v_n)^2 \\ &= (u_1^2 - Dv_1^2)(u_n^2 - Dv_n^2) = 1. \end{aligned}$$

Dunque (u_{n+1}, v_{n+1}) è soluzione per l'equazione (2.6). Osserviamo che $\forall n \in \mathbb{N}$, vale

$$u_n + v_n\sqrt{D} = (u_{n-1} + v_{n-1}\sqrt{D})(u_1 + v_1\sqrt{D}) = \cdots = (u_1 + v_1\sqrt{D})^n. \tag{2.8}$$

Denotiamo quindi con $z_n := (u_1+v_1\sqrt{D})^n,\ n \geq 0$ e notiamo che $z_0 < z_1 < \cdots < z_n < \ldots$. Dimostriamo ora che tutte le soluzioni dell'equazione (2.6) soddisfano la relazione (2.8).

Poniamo per assurdo che (u, v) sia una soluzione dell'equazione (2.6) tale che $z = u + v\sqrt{D}$ non sia nella forma della relazione (2.8). Allora $\exists\ m \in \mathbb{Z}$ t.c. $z_m < z < z_{m+1}$. Segue:

$$1 < (u + v\sqrt{D})(u_m - v_m\sqrt{D}) < u_1 + v_1\sqrt{D},$$

inoltre

$$(uu_m - Dvv_m) + (u_mv - uv_m)\sqrt{D} < u_1 + v_1\sqrt{D}.$$

D'altra parte vale che

$$(uu_m - Dvv_m)^2 - D(u_mv - uv_m)^2 = (u^2 - Dv^2)(u_m^2 - Dv_m^2) = 1,$$

per cui la coppia $(uu_m - Dvv_m,\ u_mv - uv_m)$ è soluzione della (2.6) più piccola di (u_1, v_1), il che contraddice l'ipotesi di minimalità di v_1. Ciò conclude la dimostrazione. □

Vediamo ora un'utile applicazione delle equazioni di Pell, ovvero come ricondurre la ricerca di soluzioni intere dell'equazione di una conica cartesiana alle equazioni di Pell.

Esempio 2.2 Sia data la generica equazione di una conica nel piano cartesiano

$$ax^2 + bxy + cy^2 + dx + ey + f = 0 \tag{2.9}$$

a coefficienti $a, b, c, d, e, f \in \mathbb{Z}$. Troviamo le soluzioni intere riconducendoci alle soluzioni delle equazioni di Pell.

Introduciamo il **discriminante** dell'equazione (2.9) come $\Delta = b^2 - 4ac$.

Quando $\Delta < 0$ la conica è un'ellisse, per cui posso avere al più un numero finito di soluzioni. Se invece $\Delta = 0$ la conica risulta essere una parabola. In questo caso, possiamo considerare l'equazione quando $2ae - bd = 0$ e quando $2ae - bd \neq 0$.
- **$2ae - bd = 0$**) Dalla condizione sul discriminante otteniamo $c = \frac{b^2}{4a}$

$$ax^2 + bxy + cy^2 + dx + ey + f = 0 \iff$$

$$4a^2x^2 + 4abxy + b^2y^2 + 4a(dx + ey + f) = 0 \iff$$

$$4a^2x^2 + 4abxy + b^2y^2 + 4adx + 2bdy + 4af = 0 \iff$$

$$(2ax + by + d)^2 = d^2 - 4af \iff$$

$$2ax + by = \pm\sqrt{d^2 - 4af} - d$$

Notiamo che separando i due casi (il caso con $+\sqrt{d^2 - 4af}$ e quello con $-\sqrt{d^2 - 4af}$), se $\sqrt{d^2 - 4af} \in \mathbb{Z}$ otteniamo due equazioni diofantee lineari nelle variabili x, y che sappiamo come risolvere.

- **$2ae - bd \neq 0$**) come nel caso precedente, vale la condizione sul discriminante. Applichiamo poi la sostituzione

$$X = 2ax + by + d \quad \text{e} \quad Y = (4ae - 2bd)y + 4af - d^2,$$

da cui otteniamo che l'equazione (2.9) si riduce all'equazione $X^2 + Y = 0$ che è facilmente risolvibile.

Consideriamo infine il caso $\Delta > 0$, per cui la conica (2.9) è un'iperbole. In questo caso, è possibile ricondurre l'iperbole ad un'equazione nella forma di Pell. Ne vediamo due esempi semplici espliciti per dare un'idea del procedimento da seguire. Consideriamo $\mathcal{C}_1$ e $\mathcal{C}_2$ coniche di equazioni

$$\mathcal{C}_1 : 2x^2 - 6xy + 3y^2 = -1, \quad \mathcal{C}_2 : 3x^2 - 2y^2 = 1.$$

Le riscriviamo nei seguenti modi

$$\mathcal{C}_1 : x^2 - 3(y - x)^2 = 1, \quad \mathcal{C}_2 : (3x + 2y)^2 - 6(x + y)^2 = 1.$$

Ora per $\mathcal{C}_1$ applichiamo la sostituzione $X = x,\ Y = y - x$ e otteniamo

$$\mathcal{C}_1 : X^2 - 3Y^2 = 1,$$

mentre per $\mathcal{C}_2$ applichiamo $X = 3x + 2y,\ Y = x + y$ e otteniamo

$$X^2 - 6Y^2 = 1.$$

Un altro vantaggio che traiamo dallo studio delle equazioni di Pell è, come già anticipato, quello di ottenere una buona approssimazione per la radice quadrata di interi positivi che non sono quadrati perfetti, come rapporto di interi positivi. Infatti, se (u_n, v_n) sono soluzioni dell'equazione (2.6), allora

$$u_n - v_n\sqrt{D} = \frac{1}{u_n + v_n\sqrt{D}}$$

e quindi

$$\frac{u_n}{v_n} - \sqrt{D} = \frac{1}{v_n(u_n + v_n\sqrt{D})} < \frac{1}{v_n^2\sqrt{D}} < \frac{1}{v_n^2}.$$

Segue che

$$\lim_{n\to\infty} \frac{u_n}{v_n} = \sqrt{D}.$$

Quindi il rapporto u_n/v_n approssima $\sqrt{D}$ con un errore più piccolo di $1/v_n^2$.

Vediamo ora 2 semplici esercizi in cui possono essere utilizzate le soluzioni delle equazioni di Pell per arrivare a un risultato.

Esercizio 2.3 Sia t_m l'm-esimo numero triangolare, ovvero $t_m = \sum_{i=1}^{m} i = \frac{m(m+1)}{2}$, per $m \in \mathbb{N}$. Trovare tutti i numeri triangolari che sono quadrati perfetti.
Soluzione. l'equazione $t_x = y^2,\ y \in \mathbb{N}$ è equivalente a

$$(2x+1)^2 - 8y^2 = 1$$

e quindi all'equazione di Pell

$$u^2 - 8v^2 = 1$$

la cui soluzione fondamentale è $(u_1, v_1) = (3, 1)$ e le successive saranno

$$\begin{cases} u_n = \frac{1}{2}[(3+\sqrt{8})^n + (3-\sqrt{8})^n] \\ v_n = \frac{1}{2\sqrt{8}}[(3+\sqrt{8})^n (3-\sqrt{8})^n] \end{cases}$$

per $n \geq 1$. Notiamo poi che $(3 \pm \sqrt{8}) = (\sqrt{2} \pm 1)^2$ e concludiamo calcolando

$$x_n = \frac{u_n - 1}{2} = \left[\frac{(\sqrt{2}+1)^n - (\sqrt{2}-1)^n}{2}\right]^2.$$

Esercizio 2.4 Dimostrare che esistono infinite terne di interi consecutivi i quali sono tutti e 3 scrivibili come somma di due quadrati perfetti, eventualmente nulli.
Soluzione. Se consideriamo una terna nella forma $(x^2-1, x^2, x^2+1),\ x \in \mathbb{N}$ allora è sufficiente che sia possibile riscrivere $x^2 - 1$ come somma di quadrati, essendo che $x^2 = x^2 + 0^2$ e $x^2 + 1 = x^2 + 1^2$. Se ad esempio cercassimo gli x per cui $\exists\ y \in \mathbb{N}$ t.c. $x^2 - 1 = y^2 + y^2$ e questi fossero infiniti, avremmo concluso la dimostrazione. Consideriamo quindi la successione di $(x_n, y_n)_{n \in N}$ di coppie di soluzioni dell'equazione $x^2 - 2y^2 = 1$, con soluzione base $(x_1, y_1) = (3, 2)$. Vediamo che tutte le terne della forma $(x_n^2 - 1, x_n^2, x_n^2 + 1)$ sono tali per cui tutti i 3 termini sono scrivibili come somma di quadrati perfetti, infatti $x^2 - 1 = y_n^2 + y_n^2$.

2.3 L'equazione $ax^2 - by^2 = 1$ e l'equazione negativa di Pell

Studiamo adesso l'equazione

$$ax^2 - by^2 = 1 \tag{2.10}$$

con $a, b \in \mathbb{N}$. Il discriminante dell'equazione allora sarà $\Delta = 4ab > 0$, Quindi posso ridurla a un'equazione di Pell.

Proposizione 2.4 *Se $\exists\ k \in \mathbb{Z},\ k > 1$ t.c. $ab = k^2$ allora l'equazione* (2.10) *non ha soluzioni intere positive.*

Dimostrazione.
Sia per assurdo (x, y) una soluzione della (2.10) con x, y interi positivi. Allora necessariamente a, b sono coprimi. Essendo quindi $ab = k^2$, vale $a = k_1^2,\ b = k_2^2$ per alcuni interi positivi k_1, k_2. Quindi

$$k_1^2x^2 - k_2^2y^2 = 1 \iff (k_1x - k_2y)(k_1x + k_2y) = 1.$$

Segue che

$$1 < k_1x + k_2y = k_1x - k_2y = 1$$

Il che è assurdo. □

Definizione 2.3 Definiremo l'equazione

$$u^2 - abv^2 = 1 \tag{2.11}$$

come il *risolvente di Pell* dell'equazione (2.10).

Teorema 2.5
Se l'equazione (2.10) *ammette soluzioni negli interi positivi, definiamo* (x_0, y_0) *la sua soluzione minima, tale che* $y_0 > 0$ *sia il minimo tra tutti gli* y *delle coppie di soluzioni* (x, y)*. La soluzione generale della* (2.10) *sarà allora* $(x_n, y_n)_{n\geq 0}$*, con*

$$x_n = x_0u_n + by_0v_n, \qquad y_n = x_0u_n + ay_0v_n \tag{2.12}$$

ove $(u_n, v_n)_{n\geq 0}$ *è la soluzione generale del risolvente di Pell.*

Dimostrazione.
Mostriamo che (x_n, y_n) come sopra definiti sono soluzione della (2.10), $\forall n \in \mathbb{N}$.

$$\begin{aligned} ax_n^2 - by_n^2 &= a(x_0u_n + by_0v_n)^2 - b(y_0u_n + ax_0v_n)^2 \\ &= (ax_0^2 - by_0^2)(u_n^2 - abv_n^2) = \\ &= 1 \cdot 1 = 1 \end{aligned}$$

Viceversa, sia (x, y) una soluzione dell'equazione (2.10). Allora

$$(u, v) = (ax_0x - by_0y,\ y_0x - x_0y)$$

è soluzione del risolvente di Pell. Risolvendo allora il sistema di equazioni lineari nelle incognite x e y otteniamo $x = x_0u + by_0v,\ y = y_0u + ax_0v$, dunque (x, y) è nella forma (2.12). □

Definizione 2.4 Definiamo l'equazione

$$x^2 - dy^2 = -1 \tag{2.13}$$

come **l'equazione negativa di Pell**.

Osservazione 2.3. mentre l'equazione di Pell $x^2 - dy^2 = 1$ ammette sempre soluzioni se d è un intero positivo e non è un quadrato perfetto, l'equazione negativa di Pell ammette soluzioni solo per alcuni valori di d. Notiamo infatti che è un'equazione nella forma $ax^2 - by^2 = 1$ con $a = d,\ b = 1$. Ne troveremo quindi le soluzioni come conseguenza del Teorema (2.5). Vediamone un caso particolare.

Teorema 2.6
Sia p primo. Allora l'equazione negativa di Pell

$$x^2 - py^2 = -1$$

ammette soluzioni intere positive se e solo se $p = 2$ o $p \equiv 1 \pmod 4$.

Dimostrazione.
$\Rightarrow$) (x, y) è una soluzione dell'equazione $\Rightarrow\ x^2 \equiv -1 \pmod p$. Per cui $p = 2$ o $p \equiv 1 \pmod 4$, poiché -1 è residuo quadratico modulo p primo dispari $\iff p \equiv 1 \pmod 4$.
$\Leftarrow$) Per $p = 2$, notiamo che $x = y = 1$ è soluzione. Mostriamo che per ogni primo p nella forma $p = 4t + 1$ con $t \in \mathbb{N}$ possiamo trovare delle soluzioni. Iniziamo cercando una soluzione (x_0, y_0) di interi per l'equazione di Pell corrispondente, ovvero $x_0^2 - py_0^2 = 1$. Osserviamo che x_0 deve essere necessariamente dispari. Se fosse pari infatti ne conseguirebbe che $py_0^2 \equiv y_0^2 \equiv 3 \pmod 4$, il che è assurdo perché gli unici residui quadratici modulo 4 sono 0, 1. Quindi nella relazione

$$x_0^2 - 1 = (x_0 - 1)(x_0 + 1) = py_0^2$$

i fattori $x_0 - 1$ e $x_0 + 1$ hanno massimo comune divisore 2, er cui uno dei due fattori deve essere il doppio di un quadrato perfetto, e l'altro deve essere $2p$ volte un quadrato perfetto. Denotiamoli rispettivamente con $2x^2$ e $2py^2$. Il caso $x_0 + 1 = 2x^2,\ x_0 - 1 = 2py^2$ è assurdo perché otterremmo che $x^2 - py^2 = 1$, per cui (x, y) sarebbe una soluzione dell'equazione di Pell più piccola di (x_0, y_0), che contraddice la minimalità di quest'ultima. Segue che

$$x_0 - 1 = 2x^2, \qquad x_0 + 1 = 2py^2$$

e così troviamo $x^2 - py^2 = -1$, quindi (x, y) è la soluzione cercata per l'equazione negativa di Pell. □

Parte II

Equazioni diofantee cubiche

Capitolo 3

Premesse per il Teorema di Mordell

Definizione 3.1 Consideriamo un'equazione polinomiale $f(x,y) \in \mathbb{R}[x,y]$, di grado 3. Allora l'insieme delle sue soluzioni reali rappresentano una curva detta *cubica*.

3.1 Curve algebriche piane

In questa sezione richiameremo alcune definizioni sulle curve algebriche piane e mostreremo alcuni fatti che saranno poi utili nelle sezioni successive.

Sia K un campo generico e sia $\overline{K}$ la sua chiusura algebrica.

Definizione 3.2 Consideriamo $f \in K[T] = K[T_1, \cdots, T_n]$ polinomio non costante su K. Allora l'insieme

$$V(f) := \left\{ P = \begin{pmatrix} P_1 \\ \vdots \\ P_n \end{pmatrix} \in \mathbb{A}^n_{\overline{K}} : f(P) = f(P_1, \ldots, P_n) = 0 \right\}$$

è *l'ipersuperficie associata ad f*, dove $\mathbb{A}^n_{\overline{K}}$ è lo spazio affine standard. Analogamente è definita per $g \in K[X]_h = K[x_0, \ldots, x_n]_h$ polinomio omogeneo su K

$$V(g) = \left\{ P = [P_0 : \ldots : P_n] \in \mathbb{P}^n_{\overline{K}} : g(P) = 0 \right\}$$

nel caso proiettivo, dove $\mathbb{P}^n_{\overline{K}}$ è lo spazio proiettivo standard.

Definizione 3.3 Una ipersuperficie associata a f $V(f)$ si dice *riducibile* se $\exists f_1, f_2$ tali che $V(f_1), V(f_2) \subset V(f)$ strettamente e $V(f_1) \cup V(f_2) = V(f)$. Si dice invece che è *irriducibile* se non è riducibile

Definizione 3.4 Un'ipersuperficie affine $V \subseteq \mathbb{A}^n_{\overline{K}}$ si dice *K-razionale* se $\exists f \in K[T]$ tale che $V = V(f)$. Analogamente si può dare la definizione nel caso proiettivo per un'ipersuperficie proiettiva $V \subseteq \mathbb{P}^n_{\overline{K}}$ per cui esista un polinomio omogeneo $g \in K[X]_h$ tale che $V = V(g)$.

Definizione 3.5 Il *gruppo dei divisori*, che si indica con $Div(\mathbb{A})$ di uno spazio affine $\mathbb{A}$ è il gruppo abeliano libero generato dalle ipersuperfici K-razionali irriducibili in $\mathbb{A}^n_{\overline{K}}$. Analogamente vale per $Div(\mathbb{P})$ gruppo dei divisori di uno spazio proiettivo $\mathbb{P}$ gruppo libero generato dalle ipersuperfici K-razionali irriducibili in $\mathbb{P}^n_{\overline{K}}$.

Da adesso in poi consideriamo sempre $K = \overline{K}$.

Lemma 3.1 *Sia $g \in K[x_0, \ldots, x_n]_h$ e sia $f = g^a$ il polinomio affinizzato di g con $x_0 = 1, x_i = T_i$ per $i = 1, \ldots, n$. Allora*

$$\left(\frac{\partial g}{\partial x_i}\right)^a = \frac{\partial f}{\partial T_i} \tag{3.1}$$

Dimostrazione.
Possiamo scrivere

$$f(T_1, \ldots, T_n) = \sum_{j=0}^{d} a_j(T_1, \ldots, T_n).$$

Ora valgono le seguenti uguaglianze:

$$g = \sum_{j=0}^{d} a_j(x_1, \ldots, x_n)x_0^j \quad \text{con} \quad deg(a_j(x_1, \ldots, x_n)) = d - j$$

$$\Rightarrow \frac{\partial g}{\partial x_i} = \sum_{j=0}^{d} \frac{\partial a_j(x_1, \ldots, x_n)}{\partial x_i} x_0^j, \qquad i = 1, \ldots, n$$

$$\Rightarrow \left(\frac{\partial g}{\partial x_i}\right)^a = \sum_{j=0}^{d} \frac{\partial a_j(T_1, \ldots, T_n)}{\partial T_i} = \frac{\partial f}{\partial T_i} \qquad i = 1, \ldots, n.$$

Ovvero la tesi. □

Definizione 3.6 Siano $p_1, \ldots, p_r \in \mathbb{C}[T]$ irriducibili e sia $f = p_1^{n_1} \cdot \ldots \cdot p_r^{n_r}$. Allora il *divisore di f*, che si indica con $div(f)$ è

$$div(f) = \sum_{i=1}^{r} n_i V(p_i).$$

Definizione 3.7 Sia dato un divisore D tale che

$$D = \sum_{i=1}^{k} n_i V(p_i) = \sum_{V \in X} n_V V \quad \text{con } n_V \in \mathbb{Z} \text{ quasi tutti nulli}$$

Allora *l'ordine* o *molteplicità del divisore D in V* è $ord_V(D) := n_V$.

Definizione 3.8 Un divisore D si dice *effettivo* se $\exists f \in K[T]$ tale che $D = div(f)$.

Definizione 3.9 In $Div(\mathbb{P}^2_K)$, una *curva algebrica piana proiettiva* è un divisore effettivo in $Div(\mathbb{P}^2_K)$.

Definizione 3.10 Sia $D \in Div(\mathbb{P}^n_K)$ un divisore effettivo ed $\mathbb{L}$ una sottovarietà lineare in $\mathbb{A}^n_K$, con $m := dim(\mathbb{L})$. Esprimiamo $\mathbb{L}$ con le equazioni parametriche:

$$\begin{pmatrix} x_0 \\ \vdots \\ x_n \end{pmatrix} = A \begin{pmatrix} y_0 \\ \vdots \\ y_m \end{pmatrix} \quad \text{con} \quad A \in M_{n+1,m+1}(K)$$

e diciamo che $D = div(f)$ per un qualche $f \in K[X]_h$.

Allora il *divisore d'intersezione* è:

$$D \cdot \mathbb{L} = div(f(Ay)).$$

Definizione 3.11 Dato $\mathcal{D} = div(g) \subseteq \mathbb{P}^n_K$ con $g \in K[X]_h$, il *supporto di* $\mathcal{D}$ è l'insieme dei punti:

$$Supp(\mathcal{D}) = \{P \in \mathbb{P}^n_K : g(P) = 0\}$$

Definizione 3.12 Sia $g \in K[X]_h$ e sia $\mathcal{C} = div(g) \in Div(\mathbb{P}^n_K)$. Dato un punto $P \in \mathcal{C}$ e una retta r passante per P che non sia contenuta nel supporto della curva, definiamo la *molteplicità d'intersezione di* r *con* $\mathcal{C}$ *nel punto* P come l'ordine del divisore d'intersezione $r \cdot \mathcal{C}$ in P, e si indica con $e_P(r \cdot \mathcal{C})$. Inoltre possiamo definire la *molteplicità del punto* P *in* $\mathcal{C}$ come il minimo tra i valori di $e_P(r \cdot \mathcal{C})$ al variare di r retta in $\mathbb{P}^n_K$ passante per P, e si indica con:

$$m_P(\mathcal{C}) = min\{e_P(r \cdot \mathcal{C}) : \text{ r è retta in } \mathbb{P}^n_K, \ P \in r\}$$

Definizione 3.13 Un punto $P \in Supp(\mathcal{C})$ si dice

1. *semplice* o *liscio* se $m_P(\mathcal{C}) = 1$.
2. *singolare* se $m_P(\mathcal{C}) > 1$.
3. *doppio, triplo, m-uplo* rispettivamente se $m_P(\mathcal{C}) = 2, 3, m$.

Definizione 3.14 Se $P \in Supp(\mathcal{D})$, una retta $r \subseteq \mathbb{P}^n_K$ si dice *tangente a* $\mathcal{D}$ *in* P se $e_P(r \cdot \mathcal{D}) > m_P(\mathcal{D})$. Inoltre, una sottovarietà lineare $\mathbb{L}$ passante per P si dice *tangente a* $\mathcal{D}$ *in* P se ogni retta $r \subseteq \mathbb{L}$ passante per P è tangente a $\mathcal{D}$ in P.

Teorema 3.2 (Formula di Eulero)
Se $g \in K[X]_h$ *è un polinomio di grado* $d > 0$, *allora*

$$d \cdot g(x_1, \cdots, x_n) = \sum_{i=0}^{n} x_i \cdot \left[\frac{\partial}{\partial x_i} g(x_1, \cdots, x_n)\right]$$

Dimostrazione.
Dimostriamo prima la relazione su un monomio e completiamo poi il caso polinomiale. Sia dato un monomio $f(x_1, \ldots, x_n) = x_1^{m_1} \cdot \ldots \cdot x_n^{m_n}$ tale che $m_1 + \ldots + m_n = d$. allora

$$x_i \cdot \frac{\partial}{\partial x_i} f(x_1, \ldots, x_n) = x_i \cdot \frac{\partial}{\partial x_i}(x_1^{m_1} \ldots x_n^{m_n}) = m_i x_i \cdot x_i^{m_i - 1} \prod_{\substack{j=1, \\ j \neq i}}^{n} x_j^{m_j} = m_i f(x_1, \ldots, x_n).$$

Quindi

$$\sum_{i=1}^{n} x_i \cdot \frac{\partial}{\partial x_i} f(x_1, \ldots, x_n) = \sum_{i=1}^{n} m_i f(x_1, \ldots, x_n) = f(x_1, \ldots, x_n) \sum_{i=1}^{n} m_i = d \cdot f(x_1, \ldots, x_n)$$

Infine possiamo concludere che lo stesso risultato vale anche per $g(x_1, \ldots, x_n)$, infatti se

$$g(x_1, \ldots, x_n) = \sum_{j=1}^{k} f_j(x_1, \ldots, x_n)$$

in cui gli $f_j(x_1, \ldots, x_n)$ sono tutti monomi omogenei di grado d, vale:

$$\begin{aligned}
\sum_{i=1}^{n} x_i \cdot \frac{\partial}{\partial x_i} g(x_1, \ldots, x_n) &= \sum_{i=1}^{n} x_i \cdot \frac{\partial}{\partial x_i} \sum_{j=1}^{k} f_j(x_1, \ldots, x_n) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{k} x_i \cdot \frac{\partial}{\partial x_i} f_j(x_1, \ldots, x_n) \\
&= \sum_{j=1}^{k} \sum_{i=1}^{n} x_i \cdot \frac{\partial}{\partial x_i} f_j(x_1, \ldots, x_n) \\
&= \sum_{j=1}^{k} d \cdot f_j(x_1, \ldots, x_n) = d \cdot g(x_1, \ldots, x_n)
\end{aligned}$$

□

Mostriamo ora un risultato che ci permette di verificare rapidamente se un punto è singolare o meno con l'utilizzo delle derivate.

Teorema 3.3 (Caratterizzazione differenziale di punti singolari di un divisore)
Sia $g \in K[X]_h$ e $\mathcal{D} = div(g)$, con $char(K) = 0$. Allora un punto $P \in Supp(\mathcal{D})$ è un punto singolare di $\mathcal{D} \iff \nabla g(P) = 0$.

Se poi consideriamo $f = g^a$ il polinomio g affinizzato rispetto a $x_0 = 1, x_i = T_i$ per $i = 1, \cdots, n$, sia $\overline{\mathcal{D}} = div(f)$, allora $P \in \mathbb{A}^n_K$ è un punto singolare per $\overline{\mathcal{D}} \iff f(P) = \nabla f(P) = 0$.

Dimostrazione.
sia $d := deg(g(x_0, \ldots, x_n))$. Consideriamo una generica retta r_Q passante per $P = [P_0 : \ldots : P_n]$ con $Q = [Q_0 : \ldots : Q_n]$ un punto sulla retta distinto da P. Allora

$$r_Q : [x_0 : \ldots : x_n] = y_0 [P_0 : \ldots : P_n] + y_1 [Q_0 : \ldots : Q_n] = y_0 P + y_1 Q$$

dove $y_0 + y_1 = 1, y_0, y_1 \in \mathbb{K}$. Sappiamo che $e_P(r \cdot \mathcal{D})$ è l'esponente di y_1 in $g(y_0 P + y_1 Q)$, che può essere visto come un polinomio in y_0 e y_1 omogeneo di grado d.

P sarà un punto singolare di $\mathcal{D} \iff e_p(r_Q \cdot \mathcal{D}) \geq 2$, $\forall Q \in \mathbb{P}^n_K \backslash \{P\}$, quindi ci interessano i termini di $g(y_0 P + y_1 Q)$ aventi esponente di y_1 uguale 0 o 1. In particolare

se scriviamo g come somma di monomi $f_j(x_0,\dots,x_n) = x_{0,j}^{m_{0,j}} \cdot \ldots \cdot x_{n,j}^{m_{n,j}},\ j = 1,\dots,k$ possiamo calcolare prima di tutto i termini di esponente 0 e 1 in y_1 di $f_j(y_0P + y_1Q)$. Infatti diciamo per un j qualunque, $f = f_j(x_0,\dots,x_n) = x_0^{m_0} \cdot \ldots \cdot x_n^{m_n}$ sottintendendo l'indice j e sviluppiamo:

$$f(y_0P + y_1Q) = (y_0P_0 + y_1Q_0)^{m_0} \cdot \ldots \cdot (y_0P_n + y_1Q_n)^{m_n}$$

$$= y_0^{m_0+\ldots+m_n} \prod_{i=0}^{n} P_i^{m_i} + y_0^{m_0+\cdots+m_n-1} y_1 \sum_{i=0}^{n} \left(m_i P_i^{m_i-1} Q_i \prod_{\substack{j=0 \\ j\neq i}}^{n} P_j^{m_j} \right) + \dots$$

in cui $m_0 + \cdots + m_n = d$ e gli altri termini della sommatoria che compaiono hanno grado ≥ 2 in y_1, per cui non sono utili al nostro scopo e possiamo ignorarli. Notiamo che possiamo riscrivere i primi due termini della somma in funzione di f nel seguente modo:

$$y_0^d \prod_{i=0}^{n} P_i^{m_i} = y_0^d f(P)$$

$$y_0^{d-1} y_1 \sum_{i=0}^{n} \left(m_i P_i^{m_i-1} Q_i \prod_{\substack{j=0 \\ j\neq i}}^{n} P_j^{m_j} \right) = y_0^{d-1} y_1 \sum_{i=0}^{n} \frac{\partial f}{\partial x_i}(P) \cdot Q_i.$$

Calcoliamo ora i termini che ci interessano in $g(y_0P + y_1Q)$:

$$\begin{aligned} g(y_0P + y_1Q) &= \sum_{j=1}^{k} f_j(y_0P + y_1Q) \\ &= \sum_{j=1}^{k} \left[y_0^d f_j(P) + y_0^{d-1} y_1 \sum_{i=0}^{n} \frac{\partial}{\partial x_i} f(P) \cdot Q_i \right] + \dots \\ &= y_0^d \sum_{j=1}^{k} f_j(P) + y_0^{d-1} y_1 \sum_{i=0}^{n} \left[Q_i \cdot \sum_{j=1}^{k} \frac{\partial f}{\partial x_i}(P) \right] + \dots \\ &= y_0^d g(P) + y_0^{d-1} y_1 \sum_{i=0}^{n} Q_i \cdot \frac{\partial g}{\partial x_i}(P) + \cdots \\ &= y_0^d g(P) + y_0^{d-1} y_1 \nabla g(P) \cdot Q + \dots . \end{aligned}$$

Dunque per far sì P sia un punto singolare di $\mathcal{D}$ è necessario che i primi due termini della somma siano nulli per qualunque scelta di Q. Le condizioni che vengono imposte quindi sono:

$$\begin{cases} g(P) = 0 \\ \nabla g(P) \cdot Q = 0,\ \forall Q \in \mathbb{P}_K^n \end{cases} \iff \begin{cases} g(P) = 0 \\ \nabla g(P) = 0 \end{cases}$$

Ovvero la prima parte della tesi. Quindi un punto $P \in Supp(\mathcal{D})$, ovvero tale che $g(P) = 0$ è singolare $\iff \nabla g(P) = 0$.

Inoltre, se $f = g^a$ è il polinomio g affinizzato, il punto $\overline{P} = \begin{pmatrix} P_1 \\ \vdots \\ P_n \end{pmatrix}$ è singolare per la curva $\overline{\mathcal{D}} \iff$ il punto $P = [1 : P_1 : \ldots : P_n]$ è singolare per $\mathcal{D}$. Notiamo che valgono le relazioni

$$\begin{cases} \left(\frac{\partial}{\partial x_0} g(x_0, \ldots, x_n)\right)^a = d \cdot f(T_1, \ldots, T_n) \\ \\ \left(\frac{\partial}{\partial x_i} g(x_0, \ldots, x_n)\right)^a = \frac{\partial}{\partial T_i} f(T_1, \ldots, T_n) \quad \text{per } i = 1, \ldots, n \end{cases}$$

in cui la prima relazione viene dalla formula di Eulero (Teorema (3.2)). Infatti

$$d \cdot f(P_1, \ldots, P_n) = \sum_{i=0}^{n} \frac{\partial g}{\partial x_i}(P) \cdot P_i = \frac{\partial g}{\partial x_0}(P) \cdot P_0 = \frac{\partial g}{\partial x_0}(P)$$

in cui sottointendiamo $P_0 = 1$. Quindi il punto è singolare se e solo se:

$$\begin{cases} \frac{\partial g}{\partial x_0}(P) = f(P) = 0 \\ \frac{\partial g}{\partial x_i}(P) = \frac{\partial f}{\partial T_i}(P) = 0, \quad \text{per } i = 1, \ldots, n \end{cases}$$

e dunque se e solo se $P \in Supp(\overline{\mathcal{D}})$, essendo $f(P) = 0$, e $\nabla f(P) = 0$. □

Dimostriamo infine due proposizioni, che ci saranno utili quando costruiremo la struttura di gruppo dell'insieme dei punti razionali di una curva ellittica.

Proposizione 3.4 *Sia $\mathcal{C} = div(g)$ con $g(x_0, \ldots, x_n) \in K[x_1, \ldots, x_n]_h$ e $P = [1 : 0 : \ldots : 0]$. Consideriamo $g_i \in K[x_1, \ldots, x_n]_h$ polinomi omogenei di grado i tali che possiamo riscrivere g come*

$$g(x_0, \ldots, x_n) = \sum_{i=0}^{d} x_0^{d-i} g_i(x_1, \ldots, x_n).$$

Allora

$$m_P(\mathcal{C}) = m \iff \begin{cases} g_0 = g_1 = \ldots = g_{m-1} = 0 \\ g_m \neq 0 \end{cases}$$

Inoltre il complesso tangente a $\mathcal{C}$ in P, ovvero l'insieme di rette tangenti a $\mathcal{C}$ in P è $div(g_m)$.

Dimostrazione.
Dato P, consideriamo una retta passante per P e il generico punto $Q = [0 : q_1 : \ldots : q_n]$. Allora

$$\begin{aligned} g(y_0 P + y_1 Q) = g(y_0, y_1 q_1, \ldots, y_1 q_n) &= \sum_{i=0}^{d} y_0^{d-i} g_i(y_1 q_1, \ldots, y_1 q_n) \\ &= \sum_{i=0}^{d} y_0^{d-i} y_1^i g_i(q_1, \ldots, q_n). \end{aligned}$$

Dunque $m_P(\mathcal{C}) = m$ se e solo se l'esponente più piccolo con cui compare y_1 nella sommatoria è esattamente m per qualunque scelta di Q, ovvero $g_i = 0$ per $i = 0, \ldots, m-1$ e $g_m \neq 0$.

Infine, una retta $r = P \vee Q$ è tangente in P se e solo se $e_P(r \cdot \mathcal{C}) > m$, ovvero se si annulla anche g_m nei punti di r. Dunque se e solo se $r \subseteq div(g_m)$. □

Proposizione 3.5 *Sia $g \in K[x_0, \ldots, x_n]_h$ con $char(K) = 0$, e sia $\mathcal{C} = div(g) \in Div(\mathbb{P}^n_k)$. Allora se P è un punto semplice di $\mathcal{C}$, esiste un'unica retta tangente (semplice) a $\mathcal{C}$ in P.*

Dimostrazione.
In un opportuno sistema di riferimento, il punto P ha coordinate $P = [1 : 0 : \ldots : 0]$. Scriviamo g come segue:

$$g(x_0, \ldots, x_n) = \sum_{i=0}^{n} x_0^{d-i} g_i(x_1, \ldots, x_n) \quad \text{con} \quad g_i \in K[x_1, \ldots, x_n]_h, \ deg(g_i) = i.$$

Sappiamo allora per la Proposizione (3.4) appena dimostrata che $m_P(\mathcal{C}) = 1 \Rightarrow$ il suo complesso tangente sarà $div(g_1(x_1, \ldots, x_n))$ che ha grado 1, per cui è una retta. □

3.2 Il gruppo dei punti razionali di una cubica razionale

In questa sezione studieremo l'insieme dei punti di una cubica razionale e ne descriveremo la sua struttura di gruppo secondo una specifica legge di somma, che ci permetterà poi di dare la definizione di ordine di un punto.

Definizione 3.15 Un punto (x, y) nel piano si dice **punto razionale** se $x, y \in \mathbb{Q}$.

Definizione 3.16 Una retta nel piano si dice **retta razionale** se la sua equazione può essere espressa in coefficienti razionale, ovvero se ha equazione:

$$ax + by + c = 0$$

con $a, b, c \in \mathbb{Q}$.

Notiamo subito che la retta passante per due punti razionali è razionale: Siano dati $P = (x_P, y_P)$, $Q = (x_Q, y_Q)$ due punti razionali e r la retta passante per entrambi. Se $x_P = x_Q$, la retta r avrà equazione $x = x_P$, altrimenti avrà equazione

$$y - y_P = \frac{y_Q - y_P}{x_Q - x_P} x.$$

In entrambi i casi tutti i coefficienti sono razionali, essendo P, Q razionali.

Definizione 3.17 Analogamente diciamo che una cubica è **razionale** se l'equazione che la descrive ha coefficienti razionali.

Data una cubica razionale $\Gamma \subseteq \mathbb{Q}$ possiamo ora definire un'operazione $* : \Gamma^2 \to \Gamma$ che dati due punti della cubica ne associa univocamente un terzo, nel seguente modo: dati $P = (x_P, y_P), Q = (x_Q, y_Q) \in \Gamma$ punti razionali distinti, sia r la retta che li congiunge. Allora r sarà una retta razionale e se cerchiamo le intersezioni tra r e Γ otteniamo che le soluzioni sono gli zeri di un polinomio di terzo grado in una variabile a coefficienti razionali. Per cui se 2 delle soluzioni sono razionali, lo deve essere necessariamente anche la terza. Se quindi cerchiamo le soluzioni nella variabile x otteniamo che x_P, x_Q saranno soluzioni poiché $P, Q \in r \cap \Gamma$, per cui la terza soluzione sarà $x_R \in \mathbb{Q}$. Sia allora $R := (x_R, y_R)$ il terzo punto d'intersezione tra r e Γ. Essendo r e x_R rispettivamente retta e coordinata razionale, allora anche y_R deve essere razionale. Ne segue che R è un punto razionale. Allora l'operazione $* : \Gamma^2 \to \Gamma$ associa (P, Q) a $P * Q = R$ se $P \neq Q$.

Abbiamo bisogno di definire però ancora $P * P$: consideriamo la retta r tangente a Γ in P. Se $\exists\ Q \in \Gamma \cap r,\ Q \neq P$ allora $P * P = Q$, altrimenti $P * P = P$. L'operazione $*$ così descritta è ben definita.

Lemma 3.6 *Siano $\mathcal{C}_1 = div(g_1)$ e $\mathcal{C}_2 = div(g_2)$ due cubiche con $Supp(\mathcal{C}_1) \cap Supp(\mathcal{C}_2 = \{P_1, \dots, P_9\}$ e siano sei di questi punti, diciamo $\{P_1, \dots, P_6\}$ contenuti in una conica $\mathcal{C} = s_1 + s_2 = div(g)$ ove s_1, s_2 sono due rette distinte. Allora i tre punti rimanenti $\{P_7, P_8, P_9\}$ sono allineati.*

Dimostrazione.
Ognuno dei sei punti $P_1, \dots, P_6$ deve appartenere a una di queste due rette. Supponiamo per assurdo che quattro di questi sei punti appartengano a s_1, diciamo $P_1, P_2, P_3, P_4 \in s_1$, allora $\mathcal{C}_1 \cap s_1 \supseteq \{P_1, \dots, P_4\}$. Ma allora la molteplicità d'intersezione tra $\mathcal{C}_1$ e s_1 sarebbe almeno 4, quindi $s_1 \subseteq \mathcal{C}_1$. Analogamente si ricava che $s_1 \subseteq \mathcal{C}_2$, ma allora

$$s_1 \subseteq Supp(\mathcal{C}_1) \cap Supp(\mathcal{C}_2) = \{P_1, \dots, P_9\}$$

che è assurdo. Ne segue che ci devono essere esattamente tre punti su una retta e tre sull'altra. Sia ora $\mathcal{D}$ la cubica nel fascio $div(\alpha g_1 + \beta g_2)$ passante per il punto $R := s_1 \cap s_2$. Allora

$$\mathcal{D} \cdot s_1 \supseteq \{P_4, R, P_5, P_6\} \ \Rightarrow\ s_1 \subseteq \mathcal{D}$$
$$\mathcal{D} \cdot s_2 \supseteq \{P_1, P_2, R, P_3\} \ \Rightarrow\ s_2 \subseteq \mathcal{D}.$$

Allora tutti i sei punti della conica ed R sono punti di $\mathcal{D}$, quindi $\mathcal{C} \subseteq \mathcal{D}$. Sia allora h tale che $\mathcal{D} = div(h)$ e g_3 tale che $h = g \cdot g_3$. Essendo h un polinomio di grado 3 e g un polinomio di grado 2, allora g_3 deve avere grado 1, per cui $s = div(g_3)$ è una retta, e necessariamente deve contenere i punti di $\mathcal{D}$ che non sono contenuti in $\mathcal{C}$, ovvero P_7, P_8, P_9. Segue che P_7, P_8, P_9 sono allineati. □

Proposizione 3.7 *Data una cubica razionale $\Gamma \subseteq \mathbb{Q}$ L'operazione $*$ ha le seguenti proprietà:*

a) $P * Q = Q * P$

b) $(P * Q) * Q = P$

c) $P * (O * (Q * R)) = R * (O * (Q * P))$

Dimostrazione.
a) La retta che costruiamo per trovare $P * Q$ *è la stessa che costruiamo per trovare* $Q * P$*, per cui* $P * Q = Q * P$*.*
b) Sia r *la retta passante per* P, Q*. Sia poi* $S := P * Q \in r \cap \Gamma$ *il terzo punto individuato dall'intersezione tra la retta* r *e la cubica razionale* Γ*. Costruiamo ora la retta* s *necessaria per trovare* $S * Q$*.* s *è la retta passante per* S *e* Q *e quindi coincide con* r*, per cui il terzo punto individuato dall'intersezione tra* Γ *e* $s = r$ *è esattamente* P*, ovvero* $P = S * Q = (P * Q) * Q$*.*
c) Sia $T := P * (O * (Q * R))$ *e siano date le rette*

$$r_1 = Q \vee R, \qquad r_2 = O \vee (Q * P), \qquad r_3 = P \vee (O * (Q * R)).$$

Definiamo allora la cubica $\Gamma' = r_1 + r_2 + r_3$*, e studiamo il fascio generato dalle cubiche* Γ, Γ'*. I punti base del fascio saranno*

$$Supp(\Gamma) \cap Supp(\Gamma') = \{R, Q, Q * R, O, Q * P, (Q * P) * O, P, O * (Q * R), T\}.$$

Notiamo che date le rette

$$s_1 = Q \vee P, \qquad s_2 = O \vee (Q * R),$$

allora i sei punti $\{Q, Q * R, O, Q * P, P, O * (Q * R)\}$ *sono punti della conica* $s_1 + s_2$*. Per il Lemma* (3.6) *allora, i tre punti rimanenti* $\{R, (Q * P) * O = O * (Q * P), T\}$ *sono allineati, ovvero*

$$T = R * (O * (Q * P)).$$

□

Fissiamo adesso un qualunque punto $\mathcal{O} \in \Gamma$. Definiamo un'operazione $+ : \Gamma^2 \to \Gamma$ che associa a (P, Q) il punto $P + Q = \mathcal{O} * (P * Q)$, ovvero dati P, Q due punti della cubica razionale, costruiamo la retta r passante per entrambi e identifichiamo così il punto $R := P * Q$. Dopodiché costruiamo la retta passante per $\mathcal{O}$ ed R, che avrà un'ulteriore intersezione con Γ e che sarà appunto il punto $P + Q$.

Teorema 3.8
Sia $\Gamma \subseteq Q$ *una cubica razionale con un punto* $\mathcal{O} \in \Gamma$ *fissato. Allora* $(\Gamma, +)$ *con* $+$ *definito come sopra è un gruppo abeliano con identità* $\mathcal{O}$*.*

Dimostrazione.
Commutatività: Siano $P, Q \in \Gamma$. Allora vale

$$P + Q = \mathcal{O} * (P * Q) \overset{(3.7.a)}{=} \mathcal{O} * (Q * P) = Q + P.$$

Elemento neutro : l'elemento neutro per l'operazione $+$ è $\mathcal{O}$, infatti

$$\mathcal{O} + P = \mathcal{O} * (\mathcal{O} * P) \overset{(3.7.b)}{=} P.$$

In particolare vale $\mathcal{O} + \mathcal{O} = \mathcal{O}$, per cui $\mathcal{O}$ è l'identità.
Associatività:

$$\begin{aligned} P + (Q + R) &= \mathcal{O} * (P * (\mathcal{O} * (Q * R))) \overset{(3.7.c)}{=} \\ &= \mathcal{O} * (R * (\mathcal{O} * (Q * P))) \\ &= R + (\mathcal{O} * (Q * P)) \\ &= R + (Q + P) \overset{(3.7.a)}{=} \\ &= (P + Q) + R \end{aligned}$$

Elemento inverso : l'elemento inverso $-P$ di P tale che $P + (-P) = \mathcal{O}$ è $-P = P * (\mathcal{O} * \mathcal{O})$. Infatti

$$\begin{aligned} P + (-P) &= \mathcal{O} * (P * (P * (\mathcal{O} * \mathcal{O}))) \overset{(3.7.c)}{=} \\ &= \mathcal{O} * (\mathcal{O} * (P * (\mathcal{O} * P))) \overset{(3.7.b)}{=} \\ &= \mathcal{O} * (\mathcal{O} * \mathcal{O}) \\ &= \mathcal{O} + \mathcal{O} = \mathcal{O}. \end{aligned}$$

Dunque l'insieme dei punti di una cubica razionale con l'operazione + così definita è un gruppo abeliano. □

Chiaramente la scelta di un punto $\mathcal{O}$ elemento neutro della cubica $\mathcal{C}$ è libera, infatti se scegliamo un altro punto $\mathcal{O}'$ come elemento neutro del nostro gruppo, possiamo definire un omomorfismo di gruppi, per cui la struttura del nuovo gruppo sarà identica a quella del precedente. In particolare l'applicazione

$$P \longmapsto P' = P + \mathcal{O}'$$

è un omomorfismo di $(\mathcal{C}, \mathcal{O}, +)$ in $(\mathcal{C}', \mathcal{O}', +')$ ove $+'$ è la nuova legge additiva:

$$P +' Q = P + Q - \mathcal{O}'$$

e $\mathcal{C}' = \mathcal{C}$. Infatti

$$\begin{aligned} P' +' Q' &= P' + Q' - \mathcal{O}' \\ &= (P + \mathcal{O}') + (Q + \mathcal{O}') - \mathcal{O}' \\ &= P + \mathcal{O}' + Q + (\mathcal{O}' - \mathcal{O}') \\ &= P + \mathcal{O}' + Q + \mathcal{O} \\ &= P + Q + \mathcal{O}' = (P + Q)'. \end{aligned}$$

Per dimostrare il Teorema di Mordell utilizzeremo formule esplicite per la legge additiva di punti, per cui cercheremo di semplificare il più possibile l'equazione della curva, tramite delle applicazioni che ne lasciano invariata la struttura di gruppo, ottenendo così delle leggi esplicite più semplici per la computazione della somma di punti.

Definizione 3.18 Una cubica è espressa **in forma normale di Weierstrass** se è nella forma

$$y^2 = 4x^3 - g_2x - g_3 \tag{3.2}$$

o in quella più generale

$$y^2 = x^3 + ax^2 + bx + c \tag{3.3}$$

Il nostro obiettivo è mostrare che i punti razionali di una cubica sono in corrispondenza biunivoca con i punti razionali di una cubica in forma di Weierstrass eccetto per un insieme finito di punti noti. Utilizzeremo a questo scopo una trasformazione che non manda in generale rette in rette, per cui non è banale il fatto che la struttura di gruppo venga mantenuta allo stesso modo dalla trasformazione, essendo la legge additiva basata su una costruzione di rette e intersezioni. Nonostante ciò, la trasformazione non altererà la struttura di gruppo.

Proposizione 3.9 *Il gruppo dei punti razionali di una cubica $\mathcal{C}$ è in corrispondenza biunivoca con il gruppo dei punti razionali di una cubica $\mathcal{C}'$ in forma normale di Weierstrass a meno di un sottinsieme finito di punti.*

Inoltre la struttura di gruppo viene preservata dalla trasformazione.

Dimostrazione.
Consideriamo allora la cubica $\mathcal{C}$ nel piano proiettivo, scegliendone gli assi per il riferimento in modo che la cubica sia espressa in una forma semplice. In particolare consideriamo un punto razionale $\mathcal{O} \in \mathcal{C}$ e imponiamo che la retta tangente a $\mathcal{C}$ in $\mathcal{O}$ abbia equazione $Z = 0$, ovvero che il complesso tangente di $\mathcal{O}$ sia $div(Z)$. Questa retta intersecherà la cubica in un altro punto $\mathcal{O} * \mathcal{O}$ che, a meno di scelte particolari di $\mathcal{O}$ (ovvero quando $\mathcal{O}$ è un punto di molteplicità 3), sarà distinto da $\mathcal{O}$. Imponiamo che la tangente a $\mathcal{C}$ in $\mathcal{O} * \mathcal{O}$ abbia equazione $X = 0$. Infine scegliamo una qualunque retta passante per $\mathcal{O}$ distinta dalla retta $Z = 0$ e imponiamo che abbia equazione $Y = 0$. Infine poniamo $x = \frac{X}{Z}$, $y = \frac{Y}{Z}$. L'equazione di $\mathcal{C}$ in questo riferimento diventa della forma

$$xy^2 + (ax + b)y = cx^2 + dx + e.$$

Ora moltiplicando per x e successivamente definendo $y' = xy$ (e ridefiniamo y come y') otteniamo l'equazione

$$\begin{aligned}(xy)^2 + (ax + b)xy &= cx^3 + dx^2 + ex \\ y^2 + (ax + b)y &= cx^3 + dx^2 + ex\end{aligned}$$

e infine sostituendo nuovamente $y - \frac{1}{2}(ax + b)$ al posto di y otteniamo l'equazione

$$\begin{aligned}(y - \frac{1}{2}(ax + b))^2 + (ax + b)(y - \frac{1}{2}(ax + b)) &= cx^3 + dx^2 + ex \\ y^2 - (ax + b)y + \frac{1}{4}(ax + b)^2 + (ax + b)y - \frac{1}{2}(ax + b)^2 &= cx^3 + dx^2 + ex \\ y^2 &= cx^3 + \frac{1}{4}(ax + b)^2 + dx^2 + ex\end{aligned}$$

che è nella forma

$$y^2 = \text{ cubica in } x.$$

Infine, osserviamo che la legge di gruppo viene preservata dai cambi di riferimenti proiettivi, infatti nella definizione che abbiamo dato dell'operazione $+$, vediamo che è basata su una costruzione puramente geometrica che è ben definita a prescindere dalle coordinate scelte. □

Da adesso in poi quindi considereremo le cubiche in forma di Weierstrass. Tratteremo una classe particolare di cubiche di ampio interesse, parleremo infatti delle curve ellittiche.

Definizione 3.19 Una cubica in forma normale di Weierstrass

$$y^2 = f(x) = x^3 + ax^2 + bx + c$$

è detta **curva ellittica** se le radici di $f(x)$ sono distinte.

La condizione sulle radici di f per una curva ellittica ci permette di dire che ogni punto della curva ha una e una sola tangente ben definita. Tali cubiche sono dette non singolari, essendo che non hanno punti singolari nel loro supporto. L'importanza di questa condizione è chiara dal momento che per la trattazione svolta fino ad adesso della legge additiva di punti razionali abbiamo assunto che ci fosse sempre una tangente ben definita in ogni punto, ma sappiamo che in generale questo è vero se il punto è semplice per la Proposizione (3.5). Un esempio è la cubica di equazione

$$y^2 = x^2 + x^3$$

che è singolare nel punto $(0, 0)$, avente rette tangenti di equazioni $y = x$ e $y = -x$.

Dimostriamo adesso che dal fatto che la nostra cubica abbia radici distinte possiamo dedurre che ogni punto ha una tangente semplice ben definita. Allora siano α, β, γ le radici distinte di $f(x)$, cioè

$$f(x) = (x - \alpha)(x - \beta)(x - \gamma)$$

Sia poi $F(x, y) := y^2 - f(x) = 0$ e siano le sue derivate parziali

$$\frac{\partial F}{\partial x} = -f'(x), \qquad \frac{\partial F}{\partial y} = 2y$$

Un punto (x_0, y_0) di $\mathcal{C}$ sarà singolare (per il Teorema (3.3)) se e solo se

$$F(x_0, y_0) = \frac{\partial F}{\partial x}(x_0, y_0) = \frac{\partial F}{\partial y}(x_0, y_0) = 0.$$

Mostriamo in particolare che se esiste un punto singolare, la curva non è ellittica: Cerchiamo un punto (x_0, y_0) della cubica in cui le due derivate parziali siano nulle contemporaneamente. Allora $y_0 = 0$ dovendosi annullare la derivata parziale in y, per cui $f(x_0) = y_0^2 = 0$ e anche $f'(x_0) = 0$ dovendosi annullare la derivata parziale in x, quindi x_0 è sia radice di f, sia di f'. ma

$$f'(x) = (x - \beta)(x - \gamma) + (x - \alpha)(x - \beta) + (x - \alpha)(x - \gamma)$$

e se x_0 è una radice di f, diciamo sia $x_0 = \alpha$, allora $f(\alpha) = (\alpha - \beta)(\alpha - \gamma)$, che si annulla se e solo se $\alpha = \beta$ o $\alpha = \gamma$, ovvero se la curva non è ellittica.

Viceversa se consideriamo una curva non ellittica, detto x_0 la radice doppia di f, allora il punto $(x_0, 0)$ sarà un punto singolare della curva in cui entrambe le derivate parziali si annulleranno. Quindi concludiamo che se la curva è ellittica, tutti i suoi punti sono singolari. Quindi le tangenti nei punti del supporto della curva sono uniche per la Proposizione (3.5) e di conseguenza la legge di gruppo è ben definita.

Osservazione 3.1. Il motivo per cui è di nostro interesse trattare cubiche non singolari sta nel fatto che il loro comportamento è totalmente diverso da quello delle cubiche singolari, che saranno invece studiabili con dei metodi simili a quelli delle coniche e che ne rendono la trattazione più semplice.

3.3 Formule esplicite per la legge di gruppo

Cercheremo adesso delle formule esplicite di computazione per la legge di gruppo che siano nella forma più immediata possibile e che ci aiuteranno ad arrivare a dimostrare il teorema di Mordell, nel caso di cubiche non singolari.

Consideriamo l'equazione

$$y^2 = x^3 + ax^2 + bx + c \tag{3.4}$$

e omogeneizziamola ponendo $x = \frac{X}{Z},\ y = \frac{Y}{Z}$. Così otteniamo l'equazione

$$Y^2Z = X^3 + aX^2Z + bXZ^2 + cZ^3 \tag{3.5}$$

l'intersezione della cubica con la retta all'infinito $Z = 0$ ci fornisce l'equazione $X^3 = 0$, per cui la cubica ha tre punti d'intersezione con la retta all'infinito, ma risultano essere lo stesso punto. Dunque la cubica ha esattamente 1 punto all'infinito. Sia allora $F(X, Y, Z) = Y^2Z - (X^3 + aX^2Z + bXZ^2 + cZ^3)$ e sia $\mathcal{C} = div(F)$ la cubica avente equazione (3.5). Il gradiente di F è:

$$\nabla F(X, Y, Z) = \begin{pmatrix} -3X^2 - 2aXZ - bZ^2 \\ 2YZ \\ Y^2 - (aX^2 + 2bXZ + 3cZ^2) \end{pmatrix}$$

e se lo calcoliamo nel punto all'infinito $[0 : 1 : 0]$ otteniamo:

$$\nabla F(0, 1, 0) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Per cui è un punto semplice (Teorema (3.3)). Dunque ogni cubica nella forma di Weierstrass ha uno e un solo punto all'infinito che è non singolare. Chiamiamo $\mathcal{O}$ tale punto, che è razionale e sarà l'elemento neutro della nostra legge di gruppo.

La curva che considereremo allora sarà la cubica affine nel piano xy con l'aggiunta di un punto all'infinito. Con questa definizione davvero ogni retta interseca la cubica in tre

punti. In particolare la retta all'infinito la intersecherà in $\mathcal{O}$ con molteplicità 3, le rette verticali intersecheranno la cubica due volte nel piano xy e una volta nel punto $\mathcal{O}$. Le altre rette, infine, intersecheranno la cubica in tre punti del piano xy (ammettendo negli ultimi due casi, anche le soluzioni complesse).

Con la scelta che abbiamo appena fatto per il punto $\mathcal{O}$ la costruzione del punto $P+Q$ risulterà essere geometricamente più immediata. Infatti per computare $P+Q$ prima di tutto troviamo il punto $P*Q$ costruendo la retta passante per P e Q, dopodiché costruiamo la retta per il punto $P*Q$ e il punto $\mathcal{O}$ trovando l'altro punto d'intersezione, che sarà $P+Q$. Se però il punto $\mathcal{O}$ è il punto all'infinito, la retta passante per $P*Q$ e $\mathcal{O}$ sarà la retta verticale passante per $P*Q$. Concludiamo notando che una cubica in forma di Weierstrass è simmetrica rispetto all'asse x, per cui l'ulteriore intersezione tra la retta verticale e la cubica in forma di Weierstrass, ovvero $P+Q$, sarà esattamente il punto simmetrico a $P*Q$ rispetto all'asse x.

Diventa quindi facile calcolare l'opposto di un punto $P=(x,y)$, che sarà semplicemente $-P=(x,-y)$, infatti la retta passante per P e $-P$ sarà una retta verticale e quindi intersecherà la cubica in $\mathcal{O}$, il cui simmetrico rispetto all'asse x è $\mathcal{O}$ stesso.

Osservazione 3.2. Da ciò segue anche il fatto che $P+Q+R=\mathcal{O}$ se e solo se P,Q,R sono allineati: $P*Q$ e $P+Q$ sono uno l'opposto dell'altro essendo simmetrici rispetto all'asse x per costruzione. Per cui $P+Q+(P*Q)=\mathcal{O}$. Ora se R è un punto tale per cui $P+Q+R=\mathcal{O}$, possiamo aggiungere da entrambe le parti dell'uguaglianza $P*Q=-(P+Q)$ e otteniamo

$$R=P+Q+R-(P+Q)=\mathcal{O}-(P+Q)=-(P+Q)=P*Q,$$

quindi $P+Q+R=\mathcal{O} \iff R=P*Q \iff P,Q,R$ sono tre punti allineati della cubica.

Troviamo ora delle formule esplicite per computare $P+Q$ in maniera efficiente. Siano

$$P_1=(x_1,y_1),\quad P_2=(x_2,y_2),\quad P_1*P_2=(x_3,y_3),\quad P+Q=(x_3,-y_3);$$

dati P_1,P_2 allora calcoliamo x_3,y_3. La retta passante per P_1 e P_2 ha equazione

$$y=\lambda x+\nu,\quad \text{con}\quad \lambda=\frac{y_2-y_1}{x_2-x_1}\ \text{e}\ \nu=y_1-\lambda x_1=y_2-\lambda x_2. \tag{3.6}$$

Sostituiamo l'equazione della retta in quella della cubica e otteniamo che vale la seguente uguaglianza:

$$0=x^3+(a-\lambda^2)x^2+(b-2\lambda\nu)x+(c-\nu^2).$$

che ci restituirà come soluzioni le ascisse dei tre punti d'intersezione tra la retta e la cubica, ovvero x_1,x_2,x_3. Quindi:

$$x^3+(a-\lambda^2)x^2+(b-2\lambda\nu)x+(c-\nu^2)=(x-x_1)(x-x_2)(x-x_3). \tag{3.7}$$

Uguagliando i coefficienti di x^2 dei due membri dell'uguaglianza troviamo che

$$a-\lambda^2=-x_1-x_2-x_3$$

e quindi

$$x_3 = \lambda^2 - a - x_1 - x_2 \text{ e } y_3 = \lambda x_3 + \nu.$$

Questa procedura però non è applicabile per provare a computare $P + P$, essendo che λ non può essere definito come $\lambda = \frac{y_1 - y_2}{x_1 - x_2}$. Abbiamo però definito la retta passante per P_0 e P_0 come la retta tangente alla cubica in P_0 stesso. Dalla relazione $y^2 = f(x)$ troviamo che:

$$\lambda = \left(\frac{dy}{dx}\right)_{\Big|_{P_0}} = \frac{f'(x_0)}{2y_0}.$$

Procediamo esattamente come prima per concludere il calcolo. In particolare, se $P = (x, y)$ e $2P = (x_0, y_0)$ otteniamo

$$x_0 = \lambda^2 - a - 2x = \frac{(f'(x))^2}{4y^2} - a - 2x = \frac{(3x^2 + 2ax + b)^2 - (a + 2x)(4y^2)}{4y^2}$$

e sostituendo $y^2 = f(x) = x^3 + ax^2 + bx + c$:

$$x_0 = \frac{9x^4 + 4a^2x^2 + b^2 + 12ax^3 + 4abx + 6bx^2 - 4(ax^3 - a^2x^2 - abx - ac) - 8(x^4 - ax^3 - bx^2 - cx)}{4x^3 + 4ax^2 + 4bx + 4c}.$$

Da cui

$$x_0 = \frac{x^4 - 2bx^2 - 8cx + b^2 - 4ac}{4x^3 + 4ax^2 + 4bx + 4c}. \tag{3.8}$$

Questa formula per calcolare x_0 è detta *formula di duplicazione*.

Vediamo ora un esempio in cui utilizziamo esplicitamente queste formule.

Esempio 3.1 Consideriamo la cubica

$$y^2 = x^3 + 17$$

avente due punti razionali $P_1 = (-1, 4)$ e $P_2 = (2, 5)$. Calcoliamo $P_1 + P_2$ trovando la retta passante per i due punti, di equazione

$$y = \frac{1}{3}x + \frac{13}{3}$$

per cui abbiamo $\lambda = \frac{1}{3},\ \nu = \frac{13}{3}$. Dopodiché

$$x_3 = \lambda^2 - x_1 - x_2 = -\frac{8}{9} \quad \text{e} \quad y_3 = \lambda x_3 + \nu = \frac{109}{27}$$

per cui

$$P_1 + P_2 = (x_3, -y_3) = \left(-\frac{8}{9}, -\frac{109}{27}\right).$$

Calcoliamo anche $2P_1$. Utilizziamo la formula di duplicazione per trovare le ascisse

$$x_0 = \frac{1 + 8 \cdot 17}{-4 + 17 \cdot 4} = \frac{137}{64}$$

e di conseguenza, sostituendo nell'equazione della cubica, o trovando che la retta tangente alla cubica in P_1 ha equazione $y = \frac{3}{8}x + \frac{35}{8}$ si ottiene $y_0 = -\frac{2651}{512}$. Segue che

$$2P_1 = \left(\frac{137}{64}, -\frac{2651}{512}\right).$$

3.4 Altezza di un punto

In questa sezione vediamo uno strumento molto importante per la dimostrazione del Teorema di Mordell, ovvero l'altezza di un punto razionale, che esprime da un punto di vista della teoria dei numeri quanto è complicato un punto. Denoteremo da qua in poi con $\mathcal{C}(\mathbb{Q})$ il gruppo dei punti razionali di una cubica non singolare $\mathcal{C}$.

Definizione 3.20 Sia $x = \frac{m}{n} \in \mathbb{Q}$ un numero razionale ridotto ai minimi termini. Definiamo *l'altezza di* x come il massimo tra $|m|$ e $|n|$ e la denotiamo con

$$H(x) = H\left(\frac{m}{n}\right) = max\{|m|, |n|\}.$$

L'altezza di un numero razionale è sempre un numero positivo. Questa definizione ci permette di distinguere in maniera evidente anche numeri razionali molto simili in termini di valore assoluto. Per esempio $\frac{1}{2}$ e $\frac{99999}{200000}$ hanno quasi lo stesso valore assoluto, ma la loro altezza è molto diversa ed è chiaro che il secondo numero sia più complicato del primo da trattare. Una proprietà che rende più chiara l'utilità dello strumento che stiamo utilizzando è la seguente:

Proposizione 3.10 (*Proprietà di finitezza dell'altezza*) *L'insieme di tutti i numeri razionali la cui altezza è minore o uguale di un valore fissato è un insieme finito.*

Dimostrazione .
Sia $k > 0$ fissato (scegliamo $k \in \mathbb{N}$). L'insieme dei punti razionali con altezza minore o uguale a k deve essere un sottoinsieme di $\{x = \frac{m}{n} : m \geq k,\ n \geq k\}$ necessariamente, che è un insieme finito, da cui la tesi. □

Definizione 3.21 Se $\mathcal{C}$ è una cubica non singolare

$$\mathcal{C}: \ y^2 = f(x) = x^3 + ax^2 + bx + c$$

con $a, b, c \in \mathbb{Z}$, se $P = (x, y)$ è un punto di $\mathcal{C}$, allora *l'altezza di* P è l'altezza delle sue ascisse, ovvero

$$H(P) = H(x)$$

e poniamo per convenzione $H(\mathcal{O}) = 1$.

Definizione 3.22 L'altezza h *di* P è

$$h(P) = logH(P)$$

ed è sempre un numero reale non-negativo.

Enunceremo adesso quattro lemmi che ci permetteranno di dimostrare il Teorema di Mordell. Il nostro obiettivo per il resto del capitolo sarà riuscire a dimostrarli.

Lemma 3.11 *per ogni numero reale M, l'insieme*

$$\{P \in \mathcal{C}(\mathbb{Q}) : h(P) \leq M\}$$

è finito.

Dimostrazione.
$h(P) \leq M \iff H(P) \leq e^M$. Dunque

$$\{P \in \mathcal{C}(\mathbb{Q}) : h(P) \leq M\} = \{P \in \mathcal{C}(\mathbb{Q}) : H(P) \leq e^M\}$$

che è un insieme finito per la proprietà di finitezza dell'altezza. □

I prossimi tre lemmi richiedono invece argomenti più complessi, per cui li enunciamo insieme e li riprendiamo poi uno alla volta per arrivare a dimostrarli.

Lemma 3.12 *Sia P_0 un punto razionale di $\mathcal{C}$ fissato. Allora $\exists$ k_0 costante (dipendente da P_0 e da $\mathcal{C}$) tale che*

$$h(P + P_0) \leq 2h(P) + k_0 \quad \forall P \in \mathcal{C}(\mathbb{Q}).$$

Lemma 3.13 *Esiste una costante k che dipende dalla curva $\mathcal{C}$ tale che*

$$h(2P) \geq 4h(P) - k \quad \forall P \in \mathcal{C}(\mathbb{Q}).$$

Lemma 3.14 *L'indice $(\mathcal{C}(\mathbb{Q}) : 2\mathcal{C}(\mathbb{Q}))$ è finito.*

3.4.1 L'altezza di $\mathbf{P + P_0}$

In questa sezione dimostreremo il *Lemma* (3.12).

Osservazione 3.3. Dato un punto $P = (x, y) \in \mathcal{C}(\mathbb{Q})$, possiamo trovare dei numeri interi m, n, e, con $e > 0$ tali che x e y si possono riscrivere ridotti ai minimi termini come

$$x = \frac{m}{e^2} \quad y = \frac{n}{e^3}.$$

Infatti se scriviamo

$$x = \frac{m}{M} \quad y = \frac{n}{N}$$

ridotti ai minimi termini con $M, N > 0$ e sostituiamo i valori nell'equazione di $\mathcal{C}$ otteniamo

$$\frac{n^2}{N^2} = \frac{m^3}{M^3} + a\frac{m^2}{M^2} + b\frac{m}{M} + c$$

$$M^3n^2 = N^2m^3 + aN^2Mm^2 + bN^2M^2m + cN^2M^3 \tag{3.9}$$

da cui concludiamo che $N^2|M^3n^2$ essendo un fattore comune di tutti i termini del membro di destra, ma $MCD(n, N) = 1 \Rightarrow N^2|M^3$. Inoltre notiamo che $M|N^2m^3$, ma $MCD(m, M) = 1 \Rightarrow M|N^2$. Utilizzando questo fatto nell'equazione (3.9) otteniamo che $M^2|N^2m^3 \Rightarrow M|N$ e riutilizzando questo fatto un'altra volta concludiamo che $M^3|N^2m^3 \Rightarrow M^3|N^2$, quindi $M^3 = N^2$.

Poniamo adesso $e = \frac{N}{M}$ (che è un intero positivo, perché sappiamo che $M|N$). Allora

$$e^2 = \frac{N^2}{M^2} = \frac{M^3}{M^2} = M \quad \text{e} \quad e^3 = \frac{N^3}{M^3} = \frac{N^3}{N^2} = N$$

da cui la tesi.

Osservazione 3.4. Dato un punto $P \in \mathcal{C}(\mathbb{Q})$ scritto nella forma $P = (m/e^2, n/e^3)$ ridotto ai minimi termini, abbiamo definito l'altezza di P rispetto alla sua coordinata x, in questo caso come il massimo tra $|m|$ ed e^2. Per cui valgono le limitazioni

$$|m| \leq H(P) \quad \text{e} \quad e^2 \leq H(P).$$

Notiamo però che possiamo allo stesso modo limitare n in funzione di $H(P)$. Infatti affermiamo che esiste una costante $K > 0$ dipendente da $\mathcal{C}$ tale che

$$|n| \leq K(H(P))^{\frac{3}{2}} \quad \forall P = \left(\frac{m}{e^2}, \frac{n}{e^3}\right) \in \mathcal{C}(\mathbb{Q}).$$

Infatti se sostituiamo le coordinate del punto P nell'equazione di $\mathcal{C}$ moltiplicata per e^6 otteniamo

$$n^2 = m^3 + ae^2m^2 + be^4m + ce^6.$$

Poi, applicando la disuguaglianza triangolare al valore assoluto di entrambi i membri dell'equazione concludiamo che

$$|n^2| \leq |m^3| + |ae^2m^2| + |be^4m| + |ce^6| \leq H(P)^3 + |a|H(P)^3 + |b|H(P)^3 + |c|H(P)^3.$$

Se scegliamo quindi $K = \sqrt{1 + |a| + |b| + |c|}$ vale la disuguaglianza $|n| \leq K(H(P))^{\frac{3}{2}}$.

Procediamo ora a dimostrare il Lemma (3.12).

Lemma 3.12 *Sia P_0 un punto razionale di $\mathcal{C}$ fissato. Allora $\exists\ k_0$ costante (dipendente da P_0 e da $\mathcal{C}$) tale che*

$$h(P + P_0) \leq 2h(P) + k_0 \quad \forall P \in \mathcal{C}(\mathbb{Q}).$$

Dimostrazione.
Se $P_0 = \mathcal{O}$ la disuguaglianza è chiaramente valida, essendo

$$h(P + \mathcal{O}) = h(P) \leq 2h(P) \leq 2h(P) + k_0$$

per ogni scelta di $k_0 \geq 0$. Consideriamo quindi ora $P_0 = (x_0, y_0) \neq \mathcal{O}$. Mostriamo che esiste un k_0 che soddisfa la disuguaglianza per tutti i punti P, eccetto per un sottinsieme finito. Infatti per ogni sottinsieme finito di punti P, consiederiamo le differenze $h(P+P_0)-2h(P)$ e scegliamo k_0 maggiore di tutti questi valori. Dimostriamo quindi il Lemma per tutti i punti $P \notin \{P_0, -P_0, \mathcal{O}\}$.

Sia $P = (x, y)$ e sia $P + P_0 = (\alpha, \beta)$. Per calcolare $H(P + P_0)$ è necessario calcolare $H(\alpha)$, per cui scriviamo α in funzione di (x, y) e (x_0, y_0). Dall'equazione 3.6 troviamo

$$\alpha + x + x_0 = \lambda^2 - a \quad \text{con} \quad \lambda = \frac{y - y_0}{x - x_0}$$

e riscrivendo otteniamo

$$\alpha = \frac{(y - y_0)^2}{(x - x_0)^2} - a - x - x_0 = \frac{(y - y_0)^2 - (x - x_0)^2(x + x_0 - a)}{(x - x_0)^2}$$

Notiamo che sviluppando tutti i calcoli al numeratore compare il termine $y^2 - x^3$ che possiamo sostituire con $ax^2 + bx + c$, ottenendo così un'espressione della forma

$$\alpha = \frac{Ay + Bx^2 + Cx + D}{Ex^2 + Fx + G}$$

in cui A, B, C, D, E, F, G sono numeri razionali espressi in termini di a, b, c e (x_0, y_0). Inoltre, se moltiplichiamo numeratore e denominatore dell'espressione per il minimo comune denominatore di A, B, C, D, E, F e G otteniamo la stessa espressione ma con dei coefficienti interi, per cui possiamo assumere A, B, C, D, E, F, G interi.

Sostituiamo ora $x = m/e^2,\ y = n/e^3$ secondo l'Osservazione (3.3) e moltiplichiamo l'espressione per e^4/e^4:

$$\alpha = \frac{Ane + Bm^2 + Cme^2 + De^4}{Em^2 + Fme^2 + Ge^4}$$

L'espressione scritta in questo modo potrebbe non essere ridotta ai minimi termini, ma la riduzione può solo far diminuire l'altezza, per cui $H(\alpha)$ è limitata dall'altezza del numero scritto nella forma appena vista, ovvero

$$H(\alpha) \leq max\left\{Ane + Bm^2 + Cme^2 + De^4|,\ |Em^2 + Fme^2 + Ge^4|\right\}$$

Ora grazie all'Osservazione (3.4) sappiamo che

$$e \leq H(P)^{\frac{1}{2}}, \quad n \leq KH(P)^{\frac{3}{2}} \ \text{ e } \ m \leq H(P),$$

con K che dipende solamente da a, b e c. Grazie a queste disuguaglianze e alla disuguaglianza triangolare riusciamo a stimare il numeratore e il denominatore della frazione, infatti

$$\begin{aligned} |Ane + Bm^2 + Cme^2 + De^4| &\leq |Ane| + |Bm^2| + |Cme^2| + |De^4| \\ &\leq (|AK| + |B| + |C| + |D|)H(P)^2 \end{aligned}$$

e

$$\begin{aligned} |Em^2 + Fme^2 + Ge^4| &\leq |Em^2| + |Fme^2| + |Ge^4| \\ &\leq (|E| + |F| + |G|)H(P)^2. \end{aligned}$$

Dunque

$$H(P + P_0) = H(\alpha) = max\{|AK| + |B| + |C| + |D|,\ |E| + |F| + |G|\}H(P)^2.$$

prendendo il logaritmo di entrambi i lati della disuguaglianza otteniamo

$$h(P + P_0) \leq 2h(P) + k_0$$

in cui

$$k_0 = log\left(max\left\{|AK| + |B| + |C| + |D|,\ |E| + |F| + |G|\right\}H(P)^2\right)$$

che dipende solamente da a, b, c, x_0 e y_0. Ciò è sufficiente per conclcudere la dimostrazione. □

3.4.2 L'altezza di 2P

In questa sezione dimostreremo il Lemma (3.13), per cui avremo bisogno di enunciare e dimostrare prima un ulteriore lemma.

Lemma 3.15 *Siano $\phi(X)$ e $\psi(X)$ polinomi a coefficienti interi senza radici complesse comuni. Sia d il più grande tra il grado di ψ e quello di ϕ. Allora:*

a) Esiste un intero $R \geq 1$, che dipende da ϕ e da ψ tale che per ogni numero razionale m/n vale

$$MCD\left(n^d\phi\left(\frac{m}{n}\right),\ n^d\psi\left(\frac{m}{n}\right)\right) \quad \textit{divide} \quad R.$$

b) Esistono delle costanti k_1 e k_2, dipendenti da ϕ e ψ tali che per ogni numero razionale m/n che non sono radici di ψ vale

$$dh\left(\frac{m}{n}\right) - k_1 \leq h\left(\frac{\phi\left(\frac{m}{n}\right)}{\psi\left(\frac{m}{n}\right)}\right) \leq dh\left(\frac{m}{n}\right) + k_2.$$

Dimostrazione.
a) Osserviamo innanzitutto che essendo che ϕ e ψ hanno grado minore o uguale a d, le quantità $n^d\phi(m/n)$ e $n^d\psi(m/n)$ sono entrambe intere, dunque è ben definito il massimo comune divisore dei due valori. Notiamo poi che ϕ e ψ sono intercambiabili, per cui assumiamo senza perdità di generalità che $deg(\phi) = d$ e $deg(\psi) = e \leq d$. Poi riscriviamo

$$n^d\phi\left(\frac{m}{n}\right) = a_0m^d + a_1m^{d-1}n + \ldots + a_dn^d$$
$$n^d\psi\left(\frac{m}{n}\right) = b_0m^en^{d-e} + b_1m^{e-1}n^{d-e+1} + \ldots + b_en^d$$

Definiamo adesso

$$\Phi(m.n) = n^d\phi\left(\frac{m}{n}\right) \quad \text{e} \quad \Psi(m.n) = n^d\psi\left(\frac{m}{n}\right).$$

Cerchiamo di stimare $MCD(\Phi(m/n),\ \Psi(m/n))$ in modo che la stima sia indipendente da m ed n.

Essendo che $\phi(X)$ e $\psi(X)$ non hanno radici complesse in comune, sono coprimi nell'anello $\mathbb{Q}[X]$, per cui possiamo trovare due polinomi $F, G \in \mathbb{Q}[X]$ tali che

$$F(X)\phi(X) + G(X)\psi(X) = 1 \tag{3.10}$$

Sia poi $A \in \mathbb{Z}$ tale che $AF(X)$ e $AG(X)$ siano polinomi a coefficienti interi. Sia poi D il massimo tra i gradi di F e G. Notiamo che A e D non dipendono da m ed n.

Ora calcoliamo l'equazione (3.10) in $X = m/n$ e moltiplichiamo entrambi i membri dell'uguaglianza per An^{D+d}. Così otteniamo

$$\left\{n^DAF\left(\frac{m}{n}\right)\right\} \cdot \Phi(m,n) + \left\{n^DAG\left(\frac{m}{n}\right)\right\} \cdot \Psi(m,n) = An^{D+d}.$$

Sia $\gamma = \gamma(m,n) = MCD(\Phi(m/n),\ \Psi(m/n))$. Dall'equazione precedente, essendo che le quantità tra parentesi graffe sono intere, γ divide entrambi i termini della somma, per cui $\gamma | An^{D+d}$. Dimostriamo ora che è possibile ottenere un'ulteriore stima migliore di quella appena ottenuta sulla divisibilità, che non sia dipendente da n, ovvero $\gamma | Aa_0^{D+d}$, dove a_0 è il coefficiente direttore di $\phi(X)$. Essendo infatti che $\gamma | \Phi(m,n)$, sicuramente γ divide anche

$$An^{D+d-1}\Phi(m,n) = Aa_0m^dn^{D+d-1} + Aa_1m^{d-1}n^{D+d} + \ldots + Aa_dn^{D+2d-1}$$

in cui ogni termine della somma a eccezione del primo ha il fattore An^{D+d} che è multiplo di γ, per cui $\gamma \mid Aa_0m^dn^{D+d-1}$. Ma allora

$$\gamma \mid MCD(An^{D+d}, Aa_0m^dn^{D+d-1}),$$

ed essendo m, n coprimi $\gamma \mid Aa_0n^{D+d-1}$. Ora possiamo iterare il procedimento: usando il fatto che γ divide $Aa_0n^{D+d-2}\Phi(m,n)$ arriviamo a mostrare che $\gamma \mid Aa_0^2n^{D+d-2}$ e ripetiamo fino a concludere che $\gamma \mid Aa_0^{D+d}$, che conclude la dimostrazione di *a)*.
b) Dimostriamo prima di tutto la disuguaglianza con la limitazione dall'alto, ovvero mostriamo che

$$h\left(\frac{\phi\left(\frac{m}{n}\right)}{\psi\left(\frac{m}{n}\right)}\right) \leq dh\left(\frac{m}{n}\right) + k_2.$$

Riscrivendola con H la disuguaglianza diventa

$$H\left(\frac{\phi\left(\frac{m}{n}\right)}{\psi\left(\frac{m}{n}\right)}\right) \leq H\left(\frac{m}{n}\right)^d e^{k_2}. \tag{3.11}$$

Vale allora

$$\begin{aligned} H\left(\frac{\phi\left(\frac{m}{n}\right)}{\psi\left(\frac{m}{n}\right)}\right) &= H\left(\frac{n^d\phi\left(\frac{m}{n}\right)}{n^d\psi\left(\frac{m}{n}\right)}\right) = H\left(\frac{a_0m^d + \ldots + a_dm^d}{b_0m^en^{d-e} + \ldots + b_en^d}\right) \\ &= max\left\{\left|a_0m^d + \ldots + a_dm^d\right|,\ \left|b_0m^en^{d-e} + \ldots + b_en^d\right|\right\} \\ &\leq max\left\{|a_0|\cdot|m^d| + \ldots + |a_d|\cdot|m^d|,\ |b_0|\cdot|m^e|\cdot|n^{d-e}| + \ldots + |b_e|\cdot|n^d|\right\} \\ &\leq max\left\{(|a_0| + \ldots + |a_d|)\,H\left(\frac{m}{n}\right)^d,\ (|b_0| + \ldots + |b_e|)\,H\left(\frac{m}{n}\right)^d\right\}. \end{aligned}$$

Allora possiamo scegliere

$$k_2 = log\left(max\left\{|a_0| + \ldots + |a_d|,\ |b_0| + \ldots + |b_e|\right\}\right)$$

e la disuguaglianza (3.11) è valida.

Dimostriamo ora invece l'altra disuguaglianza, ovvero

$$dh\left(\frac{m}{n}\right) - k_1 \leq h\left(\frac{\phi\left(\frac{m}{n}\right)}{\psi\left(\frac{m}{n}\right)}\right).$$

Dimostriamola per tutti i valori, eccetto che per un numero finito di valori razionali, per cui sarà sufficiente poi aggiustare il valore di k_1 per ottenere che la disuguaglianza continua a valere. In particolare dimostriamolo per tutti i numeri razionali m/n che non sono radici di ϕ.

Se $r \in \mathbb{Q}\backslash\{0\}$ chiaramente $h(r) = h(1/r)$. Per simmetria, eventualmente scambiando ϕ e ψ come nel punto precedente, possiamo assumere che $deg(\phi) = d$ e $deg(\psi) = e \leq d$.

Mantenendo la notazione del punto precedente sempre, cercheremo di calcolare l'altezza di

$$\xi = \frac{\phi\left(\frac{m}{n}\right)}{\psi\left(\frac{m}{n}\right)} = \frac{\Phi(m,n)}{\Psi(m,n)}.$$

Così otteniamo un'espressione per ξ nella forma di un rapporto tra numeri interi, per cui $H(\xi)$ sarà il massimo tra $|\Phi(m,n)|$ e $|\Psi(m,n)|$ a meno di dividere per fattori comuni.

Utilizziamo il risultato ottenuto nel punto *a)*, quindi consideriamo $R \geq 1$ indipendente da m e da n tale che il massimo comune divisore tra $|\Phi(m,n)|$ e $|\Psi(m,n)|$ sia un divisore di R. Ciò rende vere le seguenti limitazioni

$$H(\xi) \geq \frac{1}{R} max\left\{|\Phi(m,n)|, |\Psi(m,n)|\right\}$$

$$= \frac{1}{R} max\left\{\left|n^d\phi\left(\frac{m}{n}\right)\right|, \left|n^d\psi\left(\frac{m}{n}\right)\right|\right\}$$

$$\geq \frac{1}{2R}\left(\left|n^d\phi\left(\frac{m}{n}\right)\right| + \left|n^d\psi\left(\frac{m}{n}\right)\right|\right)$$

in cui l'ultima disuguaglianza vale perché $max\{a,b\} \geq \frac{1}{2}(a+b)$. Consideriamo ora il quoziente

$$\frac{H(\xi)}{H(m/n)^d} \geq \frac{1}{2R} \cdot \frac{\left|n^d\phi\left(\frac{m}{n}\right)\right| + \left|n^d\psi\left(\frac{m}{n}\right)\right|}{max\{|m|^d, |n|^d\}}$$

$$= \frac{1}{2R} \cdot \frac{\left|\phi\left(\frac{m}{n}\right)\right| + \left|\psi\left(\frac{m}{n}\right)\right|}{max\{\left|\frac{m}{n}\right|^d, 1\}}$$

Consideriamo allora la funzione

$$p(t) = \frac{|\phi(t)| + |\psi(t)|}{max\{|t|^d, 1\}}.$$

Essendo che ϕ ha grado d e ψ ha grado al massimo d, $p(t)$ ha limite diverso da 0 per $|t| \to +\infty$ che vale $|a_0|$ se ψ ha grado strettamente minore di d, mentre vale $|a_0| + |b_0|$ se ψ ha grado esattamente d. In ogni caso, al di fuori di un determinato intervallo chiuso I la funzione $p(t)$ assume valori al di fuori di un intorno di 0. All'interno di I però la funzione deve essere continua e non può annullarsi, essendo che $\phi(X)$ e $\psi(X)$ non hanno radici in comune per ipotesi. Tuttavia una funzione continua su un insieme compatto ha

massimo e minimo, e ciò vale anche quindi per l'intervallo I. In particolare essendo che $p(t)$ è sempre diversa da 0, il suo valore minimo $t_0 \in I$ deve essere positivo. Ma allora $p(t)$ assume valori al di fuoir di un intorno di 0 sia al di fuori di I, sia in I stesso, dunque esiste una costante $C_1 > 0$ tale per cui $p(t) \geq C_1 \ \forall t \in \mathbb{R}$.

Abbiamo dimostrato poco fa che

$$\frac{H(\xi)}{H(m/n)^d} \geq \frac{1}{2R} \cdot p\left(\frac{m}{n}\right).$$

Usando poi il fatto che $p(t) \geq C_1$ concludiamo che

$$H(\xi) \geq \frac{C_1}{2R} \cdot H\left(\frac{m}{n}\right)^d$$

Le costanti C_1 ed R dipendono solamente da ϕ e da ψ ma non da m ed n, quindi se prendiamo i logaritmi dell'ultima disuguaglianza, troviamo la disuguaglianza desiderata

$$h(\xi) \geq dh\left(\frac{m}{n}\right) - k_1 \ \text{ con } \ k_1 = log\left(\frac{2R}{C_1}\right).$$

Ciò conclude la dimostrazione.

□

Possiamo ora procedere a dimostrare il Lemma (3.13)

Lemma 3.13 *Esiste una costante k che dipende dalla curva $\mathcal{C}$ tale che*

$$h(2P) \geq 4h(P) - k, \qquad \forall P \in \mathcal{C}(\mathbb{Q}).$$

Dimostrazione.
Come per i lemmi precedenti, dimostriamo il lemma per tutti i punti a eccezione di un numero finito, per il quale poi semplicemente aumentiamo il valore di k fino a farlo diventare più grande di $4h(P)$ per ogni punto P in questo insieme finito rimanente di punti. Consideriamo quindi l'insieme di punti P della curva tali per cui $2P = \mathcal{O}$ e dimostriamo che è finito:

un qualunque punto dell'insieme $\mathcal{A} := \{P \in \mathcal{C} : 2P = \mathcal{O}\}$ soddisferà l'equazione $P = -P$. Essendo che se $P_i = (x_i, y_i)$ allora $-P_i = (x_i, -y_i)$, la nostra condizione è equivalente a richiedere che siano punti della curva tali che $y_i = 0$. Un qualunque punto di questa forma appartiene alla retta r di equazione $y = 0$ e dunque sarà un punto all'infinito della cubica ($\mathcal{O}$) oppure sarà un punto appartenente all'intersezione tra r e $\mathcal{C}$. perché valga tale condizione, il punto deve avere ascisse che sono soluzione dell'equazione $f(x) = 0$, quindi al massimo posso avere 3 punti che la soddisfano, dunque l'insieme è finito.

Sia ora $P = (x, y)$ e sia $2P = (\xi, \eta)$. Grazie alla formula di duplicazione (equazione (3.8)) otteniamo che

$$\xi + 2x = \lambda^2 - a \ \text{ con } \ \lambda = \frac{f'(x)}{2y}.$$

Mettendo poi tutto a comune denominatore e imponendo $y^2 = f(x)$ otteniamo una formula per ξ esplicita in funzione di x:

$$\xi = \frac{f'(x)^2 - (8x + 4a)f(x)}{4f(x)} = \frac{x^4 + \cdots}{4x^3 + \cdots}.$$

Notiamo che $f(x) \neq 0$ perché abbiamo assunto $2P \neq \mathcal{O}$.

Dunque ξ è il rapporto di due polinomi in x a coefficienti interi. Essendo che $f(x)$ è una cubica non singolare per ipotesi, $f(x)$ e $f'(x)$ non hanno radici complesse in comune. Dunque anche i polinomi al numeratore e al denominatore di ξ non hanno radici comuni. Ora essendo $h(P) = h(x)$ e $h(2P) = h(\xi)$, è sufficiente dimostrare che

$$h(\xi) \geq 4h(x) - k.$$

Però questa disuguaglianza non è altro che un caso particolare di quella che abbiamo appena mostrato nel punto *b)* del Lemma (3.15), infatti se poniamo

$$\xi = \frac{\phi(x)}{\psi(x)}, \quad x = \frac{m}{n}, \quad d = 4, \quad k_1 = k$$

otteniamo

$$4h(x) - k = dh\left(\frac{m}{n}\right) - k_1 \leq h\left(\frac{\phi(m/n)}{\psi(m/n)}\right) = h(\xi)$$

da cui la tesi. □

3.5 Il sottogruppo $2\mathcal{C}(\mathbb{Q})$

Prima di iniziare il percorso utile alla dimostrazione del Lemma (3.14) abbiamo bisogno di definire ancora una quantità associata a ogni cubica: Il discriminante.

Sia data l'equazione di una cubica in forma normale di Weierstrass

$$y^2 = f(x) = x^3 + ax^2 + bx + c$$

con $a, b, c \in \mathbb{Q}$. Siano poi $X = d^2x$ e $Y = d^3y$. Riscriviamo allora l'equazione:

$$Y^2 = X^3 + d^2aX^2 + d^4bX + d^6c.$$

Scegliendo un apposito $d \in \mathbb{Z}$ possiamo quindi eliminare i denominatori di a, b, c e assumere quindi che l'equazione della cubica sia a coefficienti interi e con $f(X)$ monico.

Definizione 3.23 Sia $f(x) = x^3 + ax^2 + bx + c$ con $a, b, c \in \mathbb{Z}$, e siano $\alpha_1, \alpha_2, \alpha_3$ le sue radici. Il *discriminante di* $f(x)$ è il numero

$$D = (\alpha_1 - \alpha_2)^2(\alpha_1 - \alpha_3)^2(\alpha_2 - \alpha_3)^2$$

Osservazione 3.5. è possibile ricavare una formula per calcolare il discriminante di $f(x)$ a partire dai suoi coefficienti. Dimostreremo adesso che vale

$$D = -4ac^3 + a^2b^2 + 18abc - 4b^3 - 27c^2.$$

Sfruttiamo prima di tutto le relazioni tra le radici di f e i suoi coefficienti

$$\begin{aligned} f(x) &= (x-\alpha_1)(x-\alpha_2)(x-\alpha_3) \\ &= x^3 - (\alpha_1+\alpha_2+\alpha_3)x^2 + (\alpha_1\alpha_2+\alpha_2\alpha_3+\alpha_1\alpha_3)x - \alpha_1\alpha_2\alpha_3 \\ &= x^3 + ax^2 + bx + c \end{aligned}$$

da cui ricaviamo

$$\begin{cases} a = -(\alpha_1+\alpha_2+\alpha_3) \\ b = \alpha_1\alpha_2+\alpha_2\alpha_3+\alpha_1\alpha_3 \\ c = -\alpha_1\alpha_2\alpha_3 \end{cases}$$

Sviluppiamo ora il discriminante di $f(x)$

$$\begin{aligned} D &= (\alpha_1-\alpha_2)^2(\alpha_1-\alpha_3)^2(\alpha_2-\alpha_3)^2 \\ &= (\alpha_1^2 - 2\alpha_1\alpha_2 + \alpha_2^2)(\alpha_1^2 - 2\alpha_1\alpha_3 + \alpha_3^2)(\alpha_2^2 - 2\alpha_2\alpha_3 + \alpha_3^2) \\ &= \alpha_1^4\alpha_2^2 + \alpha_1^4\alpha_3^2 + \alpha_1^2\alpha_2^4 + \alpha_1^2\alpha_3^4 + \alpha_2^4\alpha_3^2 + \alpha_2^2\alpha_3^4 - \\ &\quad -2(\alpha_1^4\alpha_2\alpha_3 + \alpha_1\alpha_2^4\alpha_3 + \alpha_1\alpha_2\alpha_3^4) - 2(\alpha_1^3\alpha_2^2 + \alpha_1^3\alpha_3^3 + \alpha_2^3\alpha_3^3) + \\ &\quad +2(\alpha_1^3\alpha_2^2\alpha_3 + \alpha_1^2\alpha_2\alpha_3^2 + \alpha_1^2\alpha_2^3\alpha_3 + \alpha_1^2\alpha_2\alpha_3^3 + \alpha_1\alpha_2^3\alpha_3^2 + \alpha_1\alpha_2^2\alpha_3^3) - 6\alpha_1^2\alpha_2^2\alpha_3^2 \\ &= (*). \end{aligned}$$

Sviluppiamo adesso l'espressione che vogliamo dimostrare essere il discriminante, in funzione delle radici di f

$$\begin{aligned} &-4a^3c + a^2b^2 + 18abc - 4b^3 - 27c^2 \\ &= -4(\alpha_1+\alpha_2+\alpha_3)^3\alpha_1\alpha_2\alpha_3 + (\alpha_1+\alpha_2+\alpha_3)^2(\alpha_1\alpha_2+\alpha_1\alpha_3+\alpha_2\alpha_3)^2 + \\ &\quad +18(\alpha_1+\alpha_2+\alpha_3)(\alpha_1\alpha_2+\alpha_1\alpha_3+\alpha_2\alpha_3)\alpha_1\alpha_2\alpha_3 - \\ &\quad -4(\alpha_1\alpha_2+\alpha_2\alpha_3+\alpha_1\alpha_3)^3 - 27(\alpha_1\alpha_2\alpha_3)^2 = (*). \end{aligned}$$

Dunque vale

$$D = -4a^3c + a^2b^2 + 18abc - 4b^3 - 27c^2$$

e in particolare vista la sua definizione, $D \neq 0$ se e solo se le sue radici sono distinte.

In questa sezione studieremo il sottogruppo $2\mathcal{C}(\mathbb{Q})$ e in particolare dimostreremo il Lemma (3.14) che afferma che l'indice $(\mathcal{C}(\mathbb{Q}) : 2\mathcal{C}(\mathbb{Q}))$ è finito. Questo risultato vale in realtà in generale e si dimostra utilizzando degli argomenti di teoria dei numeri algebrica, ma noi lo dimostreremo con l'aggiunta di un'ipotesi che ci permetterà di trattare l'argomento con l'utilizzo degli strumenti che già abbiamo. In particolare, se

$$\mathcal{C} : y^2 = f(x) = x^3 + ax^2 + bx + c$$

è l'equazione della cubica, assumiamo come ipotesi che $f(x)$ abbia almeno una radice razionale, che ci permetterà di dire che la curva ammette almeno un punto razionale P tale che $2P = \mathcal{O}$ (P è un punto di ordine 2). Lo stesso metodo dimostrativo vale se consideriamo una radice di $f(x) = 0$ e lavoriamo nel campo generato da questa radice sul campo dei numeri razionali. Sia per tutta questa sezione x_0 la radice razionale di $f(x)$, e sia $\mathcal{C}(\mathbb{Q}) = \Gamma$.

Essendo che $f(x_0) = 0$ ed f è un polinomio monico a coefficienti interi, allora $x_0 \in \mathbb{Z}$. Con un cambio di coordinate mandiamo il punto $(x_0, 0)$ nell'origine. La nuova equazione di $\mathcal{C}$ dopo il cambio di coordinate sarà nella forma

$$\mathcal{C} : y^2 = f(x) = x^3 + ax^2 + bx$$

con $a, b \in \mathbb{Z}$. Allora il punto $T = (0, 0)$ è un punto razionale di $\mathcal{C}$ e soddisfa la relazione $2T = \mathcal{O}$.

Inoltre il discriminante calcolato tramite la formula ricavata nell'Osservazione (3.5) diventa, in questo caso:

$$D = b^2(a^2 - 4b).$$

Essendo che lavoreremo sempre con cubiche non singolari, $D \neq 0$ e quindi nè b nè $a^2 - 4b$ sono nulli.

Il nostro interesse ora sarà rivolto allo studio di $\Gamma/2\Gamma$, quindi studieremo la mappa $P \longmapsto 2P$ e mostreremo che può essere decomposta in due operazioni più semplici. In particolare costruiremo una mappa che manda la curva $\mathcal{C}$ in un'altra curva $\overline{\mathcal{C}}$ e poi un'ulteriore mappa che rimanda a curva $\overline{\mathcal{C}}$ in $\mathcal{C}$, in cui la composizione delle due sarà la mappa $P \longmapsto 2P$, da $\mathcal{C}$ in $\mathcal{C}$.

L'altra curva che considereremo è la curva $\overline{\mathcal{C}}$ di equazione:

$$\overline{\mathcal{C}} : y^2 = x^3 + \overline{a}x^2 + \overline{b}x,$$

dove

$$\overline{a} = -2a \quad \text{e} \quad \overline{b} = a^2 - 4b$$

Le due curve $\mathcal{C}$ e $\overline{\mathcal{C}}$ sono strettamente legate, e avremo modo di studiarlo. In particolare infatti possiamo ripetere il passaggio da $\mathcal{C}$ a $\overline{\mathcal{C}}$, ma partendo da $\overline{\mathcal{C}}$ e arrivando quindi a definire una curva $\overline{\overline{\mathcal{C}}}$ data dall'equazione:

$$\overline{\overline{\mathcal{C}}} : y^2 = x^3 + \overline{\overline{a}}x^2 + \overline{\overline{b}}x$$

in cui valgono le relazioni

$$\overline{\overline{a}} = -2\overline{a} = 4a \quad \text{e} \quad \overline{\overline{b}} = \overline{a}^2 - 4\overline{b} = 4a^2 - 4(a^2 - 4b) = 16b.$$

quindi la curva $\overline{\overline{\mathcal{C}}}$ ha equazione

$$\overline{\overline{\mathcal{C}}} : y^2 = x^3 + 4ax + 16bx.$$

Se quindi invece di x e y consideriamo $4x$ e $8y$ rispettivamente, dividendo l'equazione intera per 64 l'equazione di $\overline{\overline{\mathcal{C}}}$ è la stessa di $\mathcal{C}$, quindi i loro rispettivi gruppi dei punti razionali $\overline{\overline{\Gamma}}$ e Γ sono tra loro isomorfi.

Definiamo ora un omomorfismo di gruppi $\phi : \mathcal{C} \to \overline{\mathcal{C}}$ che manda i punti razionali Γ nei punti razionali Γ di $\overline{\mathcal{C}}$, e allo stesso modo definiremo una mappa $\overline{\phi} : \overline{\mathcal{C}} \to \overline{\overline{\mathcal{C}}}$. Come conseguenza dell'isomorfismo tra $\overline{\overline{\mathcal{C}}}$ e $\mathcal{C}$, la composizione $\overline{\phi} \circ \phi$ sarà un omomorfismo di gruppi da $\mathcal{C}$ in $\mathcal{C}$ che risulterà essere la mappa $P \longmapsto 2P$. Sia $P = (x, y) \in \mathcal{C}$, con $x \neq 0$. Allora $\phi(x, y) = (\overline{x}, \overline{y})$ con

$$\overline{x} = x + a + \frac{b}{x} = \frac{y^2}{x^2} \quad \text{e} \quad \overline{y} = y\left(\frac{x^2 - b}{x^2}\right).$$

Mostriamo che $(\overline{x}, \overline{y}) \in \overline{\mathcal{C}}$:

$$\begin{aligned}
\overline{x}^3 + \overline{a}\overline{x}^2 + \overline{b}\overline{x} &= \overline{x}(\overline{x}^2 - 2a\overline{x} + (a^2 - 4b)) \\
&= \frac{y^2}{x^2}\left(\frac{y^4}{x^4} - 2a\frac{y^2}{x^2} + (a^2 - 4b)\right) \\
&= \frac{y^2}{x^2}\left(\frac{(y^2 - ax^2)^2 - 4bx^4}{x^4}\right) \\
&= \frac{y^2}{x^6}((x^3 + bx)^2 - 4bx^4) \\
&= \left(\frac{y(x^2 - b)}{x^2}\right)^2 = \overline{y}^2.
\end{aligned}$$

Per cui la mappa manda effettivamente i punti di $\mathcal{C}$ nei punti di $\overline{\mathcal{C}}$ ed è ben definita. Non l'abbiamo definita solamente per i punti $T = (0, 0)$ e per $\mathcal{O}$. Poniamo allora

$$\phi(T) = \phi(\mathcal{O}) = \overline{\mathcal{O}}.$$

Enunciamo allora la seguente proposizione riguardo alla mappa che stiamo studiando per formalizzarne i risultati

Proposizione 3.16 *Siano $\mathcal{C}$ e $\overline{\mathcal{C}}$ due curve ellittiche date dalle equazioni*

$$\mathcal{C} : y^2 = x^3 + ax^2 + bx \quad e \quad \overline{\mathcal{C}} : y^2 = x^3 + \overline{a}x^2 + \overline{b}x,$$

in cui $\overline{a} = -2a$ e $\overline{b} = a^2 - 4b$. Sia poi $T = (0, 0) \in \mathcal{C}$. Allora:

a) Esiste un omomorfismo $\phi : \mathcal{C} \to \overline{\mathcal{C}}$ definito nel seguente modo:

$$\phi(P) = \begin{cases} \left(\frac{y^2}{x^2}, \frac{y(x^2-b)}{x^2}\right), & \text{se } P = (x, y) \neq \mathcal{O}, T, \\ \overline{\mathcal{O}}, & \text{se } P = \mathcal{O} \text{ o } P = T. \end{cases}$$

In cui $ker(\phi) = \{\mathcal{O}, T\}$.

b) Applicando lo stesso procedimento a $\overline{\mathcal{C}}$ otteniamo la mappa $\overline{\phi} : \overline{\mathcal{C}} \to \overline{\overline{\mathcal{C}}}$. La curva $\overline{\overline{\mathcal{C}}}$ è isomorfa a $\mathcal{C}$ tramite la mappa $(x, y) \longmapsto \left(\frac{1}{4}x, \frac{1}{8}y\right)$. Esiste quindi un isomorfismo $\psi : \overline{\mathcal{C}} \to \mathcal{C}$ definito da

$$\psi(\overline{P}) = \begin{cases} \left(\frac{\overline{y}^2}{\overline{x}^2}, \frac{\overline{y}(\overline{x}^2-\overline{b})}{\overline{x}^2}\right), & \text{se } \overline{P} = (\overline{x}, \overline{y}) \neq \overline{\mathcal{O}}, \overline{T}, \\ \mathcal{O}, & \text{se } \overline{P} = \overline{\mathcal{O}} \text{ o } \overline{P} = \overline{T}. \end{cases}$$

c) La composizione $\psi \circ \phi : \mathcal{C} \to \mathcal{C}$ *è la mappa che dato* $P \in \mathcal{C}$,

$$\psi \circ \phi(P) = 2P.$$

Dimostrazione.
a) Abbiamo già mostrato prima che la mappa ϕ è ben definita, e sapendo che è un omomorfismo diventa chiaro che il suo nucleo è dato solamente da T e $\mathcal{O}$. Rimane quindi da dimostrare che ϕ è un omomorfismo. Mostriamo $\phi(P_1+P_2) = \phi(P_1)+\phi(P_2),\ \forall P_1, P_2 \in \mathcal{C}$. Se almeno uno tra P_1 e P_2 è $\mathcal{O}$, (possiamo assumere $P_2 = \mathcal{O}$ senza perdere di generalità) allora

$$\phi(P_1 + P_2) = \phi(P_1 + \mathcal{O}) = \phi(P_1) = \phi(P_1) + \overline{\mathcal{O}} = \phi(P_1) + \phi(\mathcal{O}) = \phi(P_1) + \phi(P_2).$$

Se invece uno dei due è T, diciamo $P_2 = T$ e $P_1 = P$, dobbiamo mostrare che $\phi(P+T) = \phi(P)$. Usiamo la formula esplicita per la legge di gruppo che abbiamo ricavato nella sezione (3.3), dato $P = (x, y)$:

$$P + T = (x, y) + (0, 0) = \left(\frac{b}{x}, -\frac{by}{x^2}\right).$$

Se scriviamo

$$P + T = (x(P + T), y(P + T)) \text{ e } \phi(P + T) = (\overline{x}(P + T), \overline{y}(P + T)),$$

troviamo che

$$\overline{x}(P + T) = \left(\frac{y(P + T)}{x(P + T)}\right)^2 = \frac{(-by/x^2)^2}{(b/x)^2} = \frac{y^2}{x^2} = \overline{x}(P).$$

Analogamente calcoliamo

$$\overline{y}(P + T) = \frac{y(P + T)(x(P + T)^2 - b)}{x(P + T)^2} = \frac{(-by/x^2)((b/x)^2 - b)}{(b/x)^2} = \overline{y}(P).$$

Quindi $\phi(P + T) = \phi(P)$, se $P \neq T$. Nel caso in cui $P = T$ però abbiamo che

$$\phi(T + T) = \phi(\mathcal{O}) = \overline{\mathcal{O}} = \overline{\mathcal{O}} + \overline{\mathcal{O}} = \phi(T) + \phi(T).$$

Ora mostriamo che ϕ manda gli inversi $-P$ negli inversi $-\phi(P)$:

$$\phi(-P) = \phi(x, -y) = \left(\left(\frac{-y}{x}\right)^2, \frac{-y(x^2 - b)}{x^2}\right) = -\phi(x, y) = -\phi(P).$$

Quindi per mostrare che è un omomorfismo rimane da verificare che se $P_1 + P_2 + P_3 = \mathcal{O}$, allora $\phi(P_1) + \phi(P_2) + \phi(P_3) = \overline{\mathcal{O}}$, perché se si verifica questa condizione, allora

$$\phi(P_1 + P_2) = \phi(-P_3) = -\phi(P_3) = \phi(P_1) + \phi(P_2).$$

Abbiamo già considerato i casi in cui almeno uno dei tre punti sia T o $\mathcal{O}$.

Dalla condizione $P_1 + P_2 + P_3 = \mathcal{O}$ sappiamo che i tre punti sono collineari per l'Osservazione (3.2), quindi sia $r : y = \lambda x + \nu$ la retta per P_1, P_2, P_3 (se alcuni di questi punti coincidono, la retta sarà tangente in quei punti). Mostriamo che $\phi(P_1), \phi(P_2), \phi(P_3)$ sono l'intersezione tra una retta con $\overline{\mathcal{C}}$. Notiamo che $\nu \neq 0$, perché se $\nu = 0$ la retta intersecherebbe la curva in T, che è una caso che abbiamo già escluso. La retta che interseca $\overline{\mathcal{C}}$ ha equazione

$$\overline{r} : y = \overline{\lambda}x + \overline{\nu}, \text{ dove } \overline{\lambda} = \frac{\nu\lambda - b}{\nu} \text{ e } \overline{\nu} = \frac{\nu^2 - a\nu\lambda + b\lambda^2}{\nu}$$

Mostriamo per il punto $\phi(P_1) = \phi(x_1, y_1) = (\overline{x}_1, \overline{y}_1) \in \overline{r}$:

$$\begin{aligned}\overline{\lambda}\overline{x}_1 + \overline{\nu} &= \frac{\nu\lambda - b}{\nu}\left(\frac{y_1}{x_1}\right)^2 + \frac{\nu^2 - a\nu\lambda + b\lambda^2}{\nu} \\ &= \frac{(\nu\lambda - b)y_1^2 + (\nu^2 - a\nu\lambda + b\lambda^2)x_1^2}{\nu x_1^2} \\ &= \frac{\nu\lambda(y_1^2 - ax_1^2) - b(y_1 - \lambda x_1)(y_1 + \lambda x_1) + \nu^2 x_1^2}{\nu x_1^2}\end{aligned}$$

ed ora utilizzando le relazioni

$$y_1^2 - ax_1^2 = x_1^3 + bx_1 \text{ e } y_1 - \lambda x_1 = \nu$$

otteniamo

$$\begin{aligned}&= \frac{\lambda(x_1^3 + bx_1) - b(y_1 - \lambda x_1) + \nu x_1^2}{x_1^2} \\ &= \frac{x_1^2(\lambda x_1 + \nu) - by_1}{x_1^2} \\ &= \frac{(x_1^2 - b)y_1}{x_1^2} = \overline{y_1}.\end{aligned}$$

Analogamente si verifica per $\phi(P_2)$ e $\phi(P_3)$. La condizione di collinearità però è sufficiente se i punti $\phi(P_1), \phi(P_2), \phi(P_3)$ sono tutti distinti, ma in generale è necessario dimostrare che $\overline{x}(P_1), \overline{x}(P_2), \overline{x}(P_3)$ sono le tre radici della cubica di equazione $(\overline{\lambda}x + \overline{\nu})^2 = \overline{f}(x)$. Allora x_1, x_2, x_3 sono le radici di $(\lambda x + \nu)^2 = f(x)$ se e solo se

$$\begin{aligned}x^3 + (a - \lambda^2)x^2 + (b - 2\lambda\nu)x - \nu^2 = x^3 - (x_1 + x_2 + x_3)x^2 + \\ + (x_1x_2 + x_1x_3 + x_2x_3)x - (x_1x_2x_3)\end{aligned}$$

ossia se e solo se valgono

$$\begin{cases}\lambda^2 - a = x_1 + x_2 + x_3 \\ b - 2\lambda\nu = x_1x_2 + x_1x_3 + x_2x_3 \\ \nu^2 = x_1x_2x_3.\end{cases}$$

Se $\overline{x}_i$ sono le immagini dei punti $P_i = (x_i, y_i)$, allora valgono $\overline{x}_i = (y_i/x_i)^2$ per $i = 1, 2, 3$. Ricordiamo che

$$\left(\frac{y_i}{x_i}\right)^2 = x_i + a + \frac{b}{x_i} = \overline{x}_i$$

e calcoliamo

$$\begin{aligned}
\overline{\lambda}^2 - \overline{a} &= \lambda^2 - 2\lambda\frac{b}{\nu} + \frac{b^2}{\nu^2} + 2a \\
&= (\lambda^2 - a) + 3a + b \cdot \left(\frac{b - 2\lambda\nu}{\nu^2}\right) \\
&= x_1 + x_2 + x_3 + 3a + b \cdot \left(\frac{x_1x_2 + x_1x_3 + x_2x_3}{x_1x_2x_3}\right) \\
&= \left(x_1 + a + \frac{b}{x_1}\right) + \left(x_2 + a + \frac{b}{x_2}\right) + \left(x_3 + a + \frac{b}{x_3}\right) \\
&= \overline{x}_1 + \overline{x}_2 + \overline{x}_3.
\end{aligned}$$

Poi, per calcolare $\overline{b} - 2\overline{\lambda}\overline{\nu}$, scriviamo prima

$$\begin{aligned}
\left(\frac{y_i}{x_i}\right)^2 \cdot \left(\frac{y_j}{x_j}\right)^2 &= \left(x_i + a - \frac{b}{x_i}\right) \cdot \left(x_j + a + \frac{b}{x_j}\right) \\
&= x_ix_j + a(x_i + x_j) + a^2 + \frac{b}{x_ix_j} \cdot \left(x_i^2 + x_j^2 + b + a(x_i + x_j)\right),
\end{aligned}$$

di conseguenza

$$\sum_{\substack{i,j=1,\\ i\neq j}}^{3} \left(\frac{y_i}{x_i}\right)^2 \cdot \left(\frac{y_j}{x_j}\right)^2 = \sum_{\substack{i,j=1,\\ i\neq j}}^{3} x_ix_j + 2a\sum_{i=1}^{3} x_i + 3a^2 + \frac{b}{x_1x_2x_3}\left(\sum_{\substack{i,j=1,\\ i\neq j}}^{3} x_i^2x_j + b\sum_{i=1}^{3} x_i + 2a\sum_{\substack{i,j=1,\\ i\neq j}}^{3} x_ix_j\right).$$

Osserviamo ora che è possibile riscrivere

$$\sum_{\substack{i,j=1,\\ i\neq j}}^{3} x_i^2x_j = \left(\sum_{\substack{i,j=1,\\ i\neq j}}^{3} x_ix_j\right) \cdot \left(\sum_{i=1}^{3} x_i\right) - 3x_1x_2x_3$$

e riscriviamo l'intera uguaglianza sfruttando le relazioni trovate tra le radici x_i e i coefficienti di $(\lambda x + \nu)^2 - f(x)$

$$\begin{aligned}
\sum_{\substack{i,j=1,\\ i\neq j}}^{3} \left(\frac{y_i}{x_i}\right)^2 \cdot \left(\frac{y_j}{x_j}\right)^2 &= b - 2\lambda\nu + 2a \cdot (\lambda^2 - a) + 3a^2 + \frac{b}{\nu^2}(b - 2\lambda\nu)(\lambda^2 - a) - 3b + 2ab\frac{b - 2\lambda\nu}{\nu^2} \\
&= -2b + a^2 - 2\lambda\nu + 2a\lambda^2 - 2b\lambda\frac{a + \lambda^2}{\nu} + b^2\frac{\lambda^2 - a}{\nu^2}.
\end{aligned}$$

Utilizziamo il risultato appena ottenuto nelle seguenti uguaglianze

$$\begin{aligned}\overline{b} - 2\overline{\lambda}\overline{\nu} &= a^2 - 4b - 2\left(\lambda - \frac{b}{\nu}\right)\cdot\left(\nu - a\lambda + b\frac{\lambda^2}{\nu}\right) = \\ &= a^2 - 4b + 2b - 2\lambda\nu + 2a\lambda^2 - \frac{2b\lambda^3}{\nu} - 2ab\frac{\lambda}{\nu} + 2\frac{b^2\lambda^2}{\nu^2} \\ &= \sum_{\substack{i,j=1,\\ i\neq j}}^{3}\left(\frac{y_i}{x_i}\right)^2\cdot\left(\frac{y_j}{x_j}\right)^2 \\ &= \overline{x}_1\overline{x}_2 + \overline{x}_1\overline{x}_3 + \overline{x}_2\overline{x}_3.\end{aligned}$$

Infine vediamo che possiamo scrivere similmente ai primi due casi, sfruttando la relazione

$$\frac{y_i}{x_i} = \lambda + \frac{\nu}{x_i}$$

che

$$\overline{\nu}^2 = \overline{x}_1\overline{x}_2\overline{x}_3.$$

Da queste tre uguaglianze possiamo dedurre che $\overline{x}_1, \overline{x}_2, \overline{x}_3$ sono esattamente le radici del polinomio $(\overline{\lambda}x + \overline{\nu})^2 - \overline{f}(x)$.

b) Abbiamo già mostrato che la curva $\overline{\overline{\mathcal{C}}}$ è data dall'equazione

$$\overline{\overline{\mathcal{C}}} : y^2 = x^3 + 4ax^2 + 16bx,$$

quindi è chiaro che la mappa $(x, y) \longmapsto (x/4, y/8)$ è un isomorfismo di $\overline{\overline{\mathcal{C}}}$ in $\mathcal{C}$. Dal punto *a)* sappiamo che esiste un omomorfismo $\overline{\phi} : \overline{\mathcal{C}} \to \overline{\overline{\mathcal{C}}}$ definita similmente a ϕ ma con $\overline{a}, \overline{b}$ al posto di a, b. Allora la mappa $\psi : \overline{\mathcal{C}} \to \mathcal{C}$ è la composizione di $\overline{\phi} : \overline{\mathcal{C}} \to \overline{\overline{\mathcal{C}}}$ con l'isomorfismo $\overline{\overline{\mathcal{C}}} \to \mathcal{C}$, per cui ψ è un omomorfismo ben definito da $\overline{\mathcal{C}}$ in $\mathcal{C}$.
c) Rimane da verificare che $\psi \circ \phi$ è la mappa che moltiplica per 2 i punti della curva.

Sfruttando nuovamente le formule esplicite per la legge di gruppo descritte nella sezione (3.3) vediamo che

$$2P = 2(x, y) = \left(\frac{(x^2 - b)^2}{4y^2}, \frac{(x^2 - b)(x^4 + 2ax^3 + 6bx^2 + 2abx + b^2)}{8y^3}\right).$$

Inoltre sappiamo che

$$\phi(x, y) = \left(\frac{y^2}{x^2}, \frac{y(x^2 - b)}{x^2}\right), \quad \psi(\overline{x}, \overline{y}) = \left(\frac{\overline{y}^2}{4\overline{x}^2}, \frac{\overline{y}(\overline{x}^2 - b)}{8\overline{x}^2}\right).$$

Calcoliamo

$$\psi \circ \phi(x,y) = \psi\left(\frac{y^2}{x^2}, \frac{y(x^2-b)}{x^2}\right)$$

$$= \left(\frac{\left(\frac{y(x^2-b)}{x^2}\right)^2}{4\left(\frac{y^2}{x^2}\right)^2}, \frac{\frac{y(x^2-b)}{x^2}\left(\left(\frac{y^2}{x^2}\right)^2 - (a^2-4b)\right)}{8\left(\frac{y^2}{x^2}\right)^2}\right)$$

$$= \left(\frac{(x^2-b)^2}{4y^2}, \frac{(x^2-b)(y^4-(a^2-4b)x^4)}{8y^3x^2}\right).$$

Ora sostituendo $y^4 = x^2(x^2+ax+b)^2$ otteniamo

$$= \left(\frac{(x^2-b)^2}{4y^2}, \frac{(x^2-b)(x^2(x^2+ax+b)^2-(a^2-4b)x^4)}{8y^3x^2}\right)$$
$$= \left(\frac{(x^2-b)^2}{4y^2}, \frac{(x^2-b)(x^4+2ax^3+6bx^2+2abx+b^2)}{8y^3}\right) = 2(x,y)$$

e quindi $\psi \circ \phi(x,y) = 2(x,y)$. Essendo poi ϕ un omomorfismo, sappiamo che

$$\phi(2P) = \phi(P) + \phi(P) + 2\phi(P).$$

Abbiamo appena mostrato che $\psi \circ \phi(P) = 2P$, quindi

$$\phi \circ \psi(\phi(P)) = 2(\phi(P)).$$

Inoltre $\phi : \mathcal{C} \to \overline{\mathcal{C}}$ è una mappa suriettiva per i punti complessi, per cui dato un punto $\overline{P} \in \overline{\mathcal{C}}$ possiamo trovare un punto $P \in \mathcal{C}$ tale che $\phi(P) = \overline{P}$, per cui vale $\phi \circ \psi(\overline{P}) = 2\overline{P}$.

Osserviamo che per poter utilizzare le formule che abbiamo appena sfruttato nella dimostrazione, è necessario che entrambe le coordinate del punto $P = (x,y)$ di partenza siano non nulle. Consideriamo quindi il caso in cui una o più coordinate del punto iniziale sono nulle. Se $x = 0$ e (x,y) è un punto di $\mathcal{C}$, allora $y = 0$, ovvero $(x,y) = (0,0) = T$. Allora $\phi(x,y) = \phi(T) = \overline{\mathcal{O}}$ e

$$\psi \circ \phi(x,y) = \psi \circ \phi(T) = \psi(\overline{\mathcal{O}}) = \mathcal{O} = 2T.$$

Se invece consideriamo $y = 0$, allora

$$\phi(x,y) = \phi(x,0) = (0,0) = \overline{T}$$

e di conseguenza $\psi \circ \phi(x,y) = \psi(\overline{T}) = \mathcal{O}$. Ricordiamo che i punti con $y = 0$ sono esattamente i punti di ordine 2, per cui $2(x,0) = \mathcal{O}$, da cui abbiamo la tesi anche in caso di una o più coordinate del punto iniziale nulle □

Nella proposizione precedente abbiamo sfruttato il fatto che ϕ sia una mappa suriettiva di punti complessi e abbiamo anche visto che un punto di Γ viene mandato in un punto di $\overline{\Gamma}$ da ϕ, ma non è scontato che un punto di $\overline{\Gamma}$ sia l'immagine di un punto di Γ.

Osservazione 3.6. Consideriamo l'immagine di Γ tramite ϕ, detta $\phi(\Gamma)$ che sarà un sottogruppo di $\overline{\Gamma}$. Cerchiamo di descrivere l'immagine con tre sue caratteristiche che adesso dimostreremo:

(i) $\overline{\mathcal{O}} \in \phi(\Gamma)$.

(ii) $\overline{T} = (0,0) \in \phi(\Gamma)$ se e solo se $\overline{b} = a^2 - 4b$ è un quadrato perfetto.

(iii) Sia $\overline{P} = (\overline{x}, \overline{y}) \in \overline{\Gamma}$ con $\overline{x} \neq 0$. Allora $\overline{P} \in \phi(\Gamma)$ se e solo se $\overline{x}$ è il quadrato di un numero razionale.

Per la prima affermazione è sufficiente notare che $\phi(\mathcal{O}) = \overline{\mathcal{O}}$. Mostriamo (ii). Dalla formula di ϕ notiamo che $\overline{T} \in \phi(\Gamma)$ se e solo se esiste un punto $(x,y) \in \Gamma$ tale che $y^2/x^2 = 0$. Qui $x \neq 0$, infatti se fosse $x = 0 \Rightarrow (x,y) = T$, ma $\phi(T) = \overline{\mathcal{O}} \neq \overline{T}$. Quindi perché $\overline{T}$ appartenga all'immagine di Γ tramite ϕ abbiamo bisogno di un punto (x,y) in Γ con $x \neq 0, y = 0$, che quindi soddisferà l'equazione

$$0 = x^3 + ax^2 + bx = x(x^2 + ax + b).$$

Le soluzioni razionali non nulle di questa equazione sono le stesse dell'equazione $x^2 + ax + b = 0$, che sono

$$x_{1,2} = \frac{-a \pm \sqrt{a^2 - 4b}}{2}$$

quindi saranno razionali se e solo se $a^2 - 4b$ sarà il quadrato di un numero razionale. Tuttavia, $a, b \in \mathbb{Z} \Rightarrow a^2 - 4b \in \mathbb{Z}$, che quindi dovrà essere un quadrato perfetto.

Infine dimostriamo l'affermazione (iii). Se $(\overline{x}, \overline{y}) \in \phi(\Gamma)$ è un punto con $\overline{x} \neq 0$, ricaviamo dalla formula di ϕ che $\overline{x} = y^2/x^2$ è il quadrato di un numero razionale. Viceversa, se $\exists w \in \mathbb{Q}$ tale che $\overline{x} = w^2$, riusciamo a trovare un punto di Γ che viene mandato in $(\overline{x}, \overline{y})$:

il kernel di ϕ ha due elementi, ovvero $\mathcal{O}$ e T, di conseguenza se $(\overline{x}, \overline{y}) \in \phi(\Gamma)$ esisteranno due punti in Γ che verranno mandati in $(\overline{x}, \overline{y})$ da ϕ. Definiamo i punti

$$x_1 = \frac{1}{2}\left(w^2 - a + \frac{\overline{y}}{w}\right), \quad y_1 = x_1 w$$
$$x_2 = \frac{1}{2}\left(w^2 - a - \frac{\overline{y}}{w}\right), \quad y_2 = -x_2 w.$$

Mostriamo che $P_i = (x_i, y_i) \in \Gamma$ e $\phi(P_i) = (\overline{x}, \overline{y})$ per $i = 1, 2$. Chiaramente P_1, P_2 sono punti razionali. Inoltre vale

$$\begin{aligned} x_1 x_2 &= \frac{1}{4}\left((w^2 - a)^2 - \frac{\overline{y}^2}{w^2}\right) \\ &= \frac{1}{4}\left((\overline{x} - a)^2 - \frac{\overline{y}^2}{\overline{x}}\right) \\ &= \frac{1}{4}\left(\frac{\overline{x}^3 - 2a\overline{x}^2 + a^2\overline{x} - \overline{y}^2}{\overline{x}}\right) \end{aligned}$$

ma ricordiamo

$$\overline{y}^2 = \overline{x}^3 - 2a\overline{x}^2 + (a^2 - 4b)\overline{x}$$

per l'equazione di $\overline{\mathcal{C}}$, da cui ricaviamo che $x_1x_2 = b$. Per mostrare che P_i è un punto del supporto di $\mathcal{C}$ dobbiamo mostrare che

$$y_i^2 = x_i^3 + ax_i^2 + bx_i$$

o equivalentemente

$$\frac{y_i^2}{x_i^2} = x_i + a + \frac{b}{x_i}.$$

Sostituendo nell'equazione $b = x_1x_2$ e le espressioni per y_1, y_2 otteniamo $y_i/x_i = \pm w$, che è analogo a mostrare che

$$w^2 = x_1 + a + x_2.$$

Rimane ora da mostrare che $\phi(P_i) = (\overline{x}, \overline{y})$, ovvero

$$\frac{y_i^2}{x_i^2} = \overline{x} \text{ e } \frac{y_i(x_i^2 - b)}{x_i^2} = \overline{y}.$$

Per la prima uguaglianza ricordiamo che $y_i = \pm x_i w$ e $\overline{x} = w^2$, quindi

$$\frac{y_i^2}{x_i^2} = \frac{x_i^2 w^2}{x_i^2} = w^2 = \overline{x}.$$

Per la seconda calcoliamo separatamente:

$$\frac{y_1(x_1^2 - b)}{x_1^2} = \frac{x_1 w(x_1^2 - x_1x_2)}{x_1^2} = w(x_1 - x_2)$$

$$\frac{y_2(x_2^2 - b)}{x_2^2} = \frac{x_2 w(x_2^2 - x_1x_2)}{x_2^2} = w(x_1 - x_2),$$

e infine

$$w(x_1 - x_2) = w\left(\frac{\overline{y}}{w}\right) = \overline{y}.$$

Ciò è sufficiente a concludere la dimostrazione di (iii).

Utilizzeremo queste proprietà per dimostrare in particolare che gli indici $(\overline{\Gamma} : \phi(\Gamma))$ e $(\Gamma : \phi(\overline{\Gamma}))$ sono finiti, da cui seguirà poi il Lemma (3.14). Per farlo ci serviremo di una mappa $\alpha : \Gamma \to \mathbb{Q}^*/\mathbb{Q}^{*2}$, dove $\mathbb{Q}^*$ è il gruppo moltiplicativo dei razionali non nulli, mentre

$$\mathbb{Q}^{*2} := \left\{u^2 : u \in \mathbb{Q}^*\right\}.$$

α sarà definita come segue:

$$\begin{aligned} \alpha(\mathcal{O}) &= 1 \pmod{\mathbb{Q}^{*2}}, \\ \alpha(T) &= b \pmod{\mathbb{Q}^{*2}}, \\ \alpha(x, y) &= x \pmod{\mathbb{Q}^{*2}}, \text{ se } x \neq 0 \end{aligned}$$

Mostreremo nella seguente proposizione che α è un omomorfismo e il suo kernel è esattamente l'immagine di ψ. Inoltre, a meno di moltiplicazioni per quadrati perfetti, mostreremo che ci sono solo un numero finito di possibilità per le ascisse di un punto sulla curva, che sarà un'informazione cruciale per la dimostrazione del Lemma (3.14).

Proposizione 3.17

a) La mappa $\alpha : \Gamma \to \mathbb{Q}^*/\mathbb{Q}^{*2}$ definita come sopra è un omomorfismo

b) il kernel di α è l'immagine $\psi(\overline{\Gamma})$. Inoltre α induce un omomorfismo biettivo

$$\Gamma/\psi(\overline{\Gamma}) \hookrightarrow \mathbb{Q}^*/\mathbb{Q}^{*2}.$$

c) Siano $p_1, p_2, \ldots, p_t$ i primi distinti che dividono b. Allora l'immagine di α è contenuta nel seguente sottogruppo di $\mathbb{Q}^*/\mathbb{Q}^{*2}$:

$$\{\pm p_1^{\epsilon_1} p_2^{\epsilon_2} \cdot \ldots \cdot p_t^{\epsilon_t} : \epsilon_i \in \{0,1\}\,,\ \textit{per } i = 1, \ldots, t\}\,.$$

d) L'indice $(\Gamma : \psi(\overline{\Gamma}))$ è al massimo 2^{t+1}.

Dimostrazione.
a) Osserviamo che α manda gli inversi $-P$ negli inversi $\alpha(P)^{-1}$:

$$\alpha(-P) = \alpha(x, -y) = x = \frac{1}{x} \cdots x^2,$$

quindi

$$\alpha(-P) \equiv \frac{1}{x} = \frac{1}{\alpha(x,y)} = \alpha(P)^{-1} \pmod{\mathbb{Q}^{*2}}.$$

Per mostrare quindi che α è un omomorfismo è sufficiente mostrare che

$$P_1 + P_2 + P_3 = \mathcal{O} \ \Rightarrow\ \alpha(P_1)\alpha(P_2)\alpha(P_3) \equiv 1 \pmod{\mathbb{Q}^{*2}}.$$

Le terne di punti che sommano a $\mathcal{O}$ consistono negli insiemi dati dalle intersezioni tra la curva e una retta. Sia $y = \lambda x + \nu$ la generica retta e siano x_1, x_2, x_3 le ascisse dei tre punti individuati dall'intersezione, allora per quanto visto nella sezione (3.3), x_1, x_2, x_3 sono le soluzioni dell'equazione

$$x^3 + (a - \lambda^2)x^2 + (b - 2\lambda\nu)x + (c - \nu^2) = 0$$

per la cubica di equazione $y^2 = x^3 + ax^2 + bx + c$. Per cui

$$\begin{aligned} x_1 + x_2 + x_3 &= \lambda^2 - a \\ x_1x_2 + x_1x_3 + x_2x_3 &= b - 2\nu \\ x_1x_2x_3 &= \nu^2 - c \end{aligned}$$

ma ricordiamo che per le assunzioni fatte in precedenza, $c = 0$, da cui

$$x_1x_2x_3 = \nu^2 \in \mathbb{Q}^2.$$

Di conseguenza

$$\alpha(P_1)\alpha(P_2)\alpha(P_3) = x_1x_2x_3 = \nu^2 \equiv 1 \pmod{\mathbb{Q}^{*2}}.$$

b) Il kernel di α è necessariamente l'immagine $\psi(\overline{\Gamma})$, infatti dalla definizione di α vediamo che i suoi punti avente prima coordinata non nulla che fanno parte del suo nucleo sono quelli che hanno come prima coordinata un quadrato perfetto. Per quanto visto nell'Osservazione (3.6), tali punti appartengono all'immagine di ψ necessariamente. Ne consegue che α induce un omomorfismo biettivo da $\Gamma/\psi(\overline{\Gamma})$ in $\mathbb{Q}^*/\mathbb{Q}^{*2}$.
c) Riscriviamo i punti di Γ come $x = m/e^2$ e $y = n/e^3$ per l'Osservazione (3.3) e sostituiamole nell'equazione della curva, eliminando il denominatore:

$$n^2 = m^3 + am^2e^2 + bme^4 = m(m^2 + ame^2 + be^4).$$

Notiamo che $MCD(m, m^2 + ame^2 + be^4) = MCD(m, be^4) = MCD(m, b)$. Quindi se $MCD(m, b) = 1$ sia m sia $m^2 + ame^2 + be^4$ devono essere quadrati perfetti a meno di segno e quindi, sempre a meno di un fattore -1, $x = m/e^2$ è il quadrato di un numero razionale. Nel caso più generale in cui $d = MCD(m, b)$ eventualmente diverso da 1, ricaviamo dall'equazione

$$n^2 = dm'(d^2(m')^2 + dam'e^2 + db'e^4)$$

dove $m = m'd, b = b'd$, $m', b' \in \mathbb{Z}$ e $MCD(m', b') = 1$ e quindi

$$\left(\frac{n}{d}\right)^2 = m'(d(m')^2 + am'e^2 + b'e^4)$$

dove $n/d \in \mathbb{Z}$ e $MCD(m', d(m')^2 + am'e^2 + b'e^4) = MCD(m', b'e^4) = MCD(m', b') = 1$, per cui ci siamo ricondotti al caso in cui i due fattori sono coprimi e m' deve essere un quadrato perfetto. Di conseguenza possiamo riscrivere m come prodotto di un quadrato perfetto, moltiplicato per il massimo comune divisore tra m e b. Riscriviamo allora:

$$m = \pm k^2 \cdot p_1^{\epsilon_1} \cdot p_2^{\epsilon_2} \cdot \ldots \cdot p_t^{\epsilon_t}, \quad \epsilon_i \in \{0, 1\}$$

dove $p_1, p_2, \ldots, p_t$ sono i primi distinti che dividono b. Allora

$$\alpha(P) = x = \frac{m}{e^2} \equiv \pm p_1^{\epsilon_1} \cdot p_2^{\epsilon_2} \cdot \ldots \cdot p_t^{\epsilon_t} \pmod{Q^{*2}}$$

e quindi l'immagine di α è contenuta nell'insieme dei punti

$$\left\{\pm p_1^{\epsilon_1} p_2^{\epsilon_2} \cdot \ldots \cdot p_t^{\epsilon_t} : \epsilon_i \in \{0, 1\}, \text{ per } i = 1, \ldots, t\right\}.$$

Se $x = 0$ e quindi $m = 0$ non possiamo applicare il procedimento appena mostrato, ma dalla definizione di α sappiamo che $\alpha(T) = b \pmod{\mathbb{Q}^{*2}}$, per cui la conclusione rimane valida perché, a meno di quadrati perfetti, b può essere scritto nella forma indicata.
d) Il sottogruppo descritto nel punto *c)* ha esattamente 2^{t+1} elementi. Inoltre nel punto *b)* abbiamo mostrato che il gruppo quoziente $\Gamma/\psi(\overline{\Gamma}$ viene mandato tramite una corrispondenza biunivoca in questo sottogruppo, per cui l'indice di $\psi(\Gamma)$ in Γ è al massimo 2^{t+1}. □

Sappiamo quindi che esistono due omomorfismi $\phi : \Gamma \to \overline{\Gamma}$ e $\psi : \overline{\Gamma} \to \Gamma$ tali che le loro composizioni $\phi \circ \psi$ e $\psi \circ \phi$ sono la mappa di moltiplicazione per 2 di un punto e gli indici $(\overline{\Gamma} : \phi(\Gamma))$ e $(\Gamma : \psi(\overline{\Gamma}))$ sono finiti. Concludiamo allora la dimostrazione del Lemma (3.14) grazie al Lemma che segue.

Lemma 3.18 *Siano A, B due gruppi abeliani e siano $\phi : A \to B, \psi : B \to A$ due omomorfismi tali che*

$$\psi \circ \phi(a) = 2a \ \forall a \in A \ e \ \phi \circ \psi(b) = 2b \ \forall b \in B$$

e tali che $\phi(A)$ abbia indice finito in B e $\psi(B)$ abbia indice finito in A. Allora $2A$ ha indice finito in A e vale la disuguaglianza

$$(A : 2A) \leq (A : \psi(B))(B : \phi(A)).$$

Dimostrazione.
Avendo $\psi(B)$ indice finito in A, esistono degli elementi $a_1, \ldots, a_n$ rappresentanti per le classi laterali. Analogamente ricaviamo gli elementi $b_1, \ldots, b_m$. Mostriamo che l'insieme

$$I := \{a_i + \psi(b_j) : 1 \leq i \leq n, 1 \leq j \leq m\}$$

include al suo interno un insieme completo di rappresentanti per le classi laterali di $2A$ in A.

Sia $a \in A$. Cerchiamo di scrivere a come somma di un elemento di I e un elemento di $2A$. $a_1, \ldots, a_n$ sono rappresentanti delle classi laterali di $\psi(B)$ in A, quindi possiamo scegliere un a_i tale per cui $a - a_i \in \psi(B)$. Sia allora $b \in B$ tale che $a - a_i = \psi(b)$. Similmente troviamo b_j tale che $b - b_j \in \phi(A)$, e sia $a' \in A$ tale che $b - b_j = \phi(a')$. Allora

$$\begin{aligned} a = a_i + \psi(b) &= a_i + \psi\left(b_j + \phi(a')\right) \\ &= a_i + \psi(b_j) + \psi(\phi(a')) \\ &= \underbrace{a_i + \psi(b_j)}_{\in I} + \underbrace{2a'}_{\in A}, \end{aligned}$$

da cui la tesi. □

Con questo lemma concludiamo la dimostrazione del Lemma (3.14), per cui abbiamo tutti gli strumenti per procedere con la dimostrazione del Teorema di Mordell, che vedremo nel prossimo capitolo.

Capitolo 4

Teorema di Mordell e sue applicazioni

In questo capitolo dimostreremo il Teorema della Discesa, che avrà come diretta conseguenza il Teorema di Mordell viste le premesse svolte nel capitolo precedente. Il nome di questo teorema deriva dalla dimostrazione che segue la stessa logica del metodo della discesa infinita di Fermat. Infatti, a partire da un punto P ci ricondurremo a trovare che una stessa proprietà deve valere anche per un punto P' che è, in un certo senso, più piccolo di P. Per definire la grandezza di un punto sfrutteremo l'altezza, di cui ora conosciamo molte proprietà utili. In seguito, a partire dalla dimostrazione del Teorema della Discesa, vedremo anche come, alcune volte, sarà possibile trovare un insieme di generatori per il gruppo dei punti razionali. Bisogna notare però che il Teorema di Mordell ci assicura solamente che un tale insieme di generatori esista, ma non è ancora stato dimostrato che il procedimento che applicheremo ci permetta sempre di ottenere l'insieme cercato.

Teorema 4.1 (Teorema della Discesa)
Sia Γ *un gruppo commutativo, e sia data una funzione*

$$h : \Gamma \to [0, \infty)$$

che rispetta le seguenti proprietà:

a) Per ogni numero reale M*, l'insieme*

$$\{P \in \Gamma : h(P) \leq M\}$$

è finito.

b) Per ogni $P_0 \in \Gamma$ *esiste una costante* k_0 *tale che*

$$h(P + P_0) \leq 2h(P) + k_0, \quad \forall P \in \Gamma.$$

c) Esiste una costante k *tale che*

$$h(2P) \geq 4h(P) - k, \quad \forall P \in \Gamma.$$

Inoltre, supponiamo anche che valga

d) Il sottogruppo 2Γ *ha indice finito in* Γ.

Allora Γ *è finitamente generato.*

Dimostrazione.
Consideriamo un rappresentante per ogni classe laterale di 2Γ in Γ. Sappiamo che esistono un numero finito di classi laterali, diciamo n, e chiamiamo $Q_1, \ldots, Q_n$ i rappresentanti cercati. Allora per ogni punto $P \in \Gamma$, esiste un indice i_1 dipendente da P tale che

$$P - Q_{i_1} \in 2\Gamma$$

perché P deve appartenere almeno a una classe laterale. Sia allora $P_1 \in \Gamma$ tale che

$$P - Q_{i_1} = P_1.$$

Ripetiamo poi lo stesso procedimento a partire da P_1 con l'indice i_2 e iteriamo ulteriormente. Scriviamo allora

$$\begin{aligned} P_1 - Q_{i_2} &= 2P_2 \\ P_2 - Q_{i_3} &= 2P_3 \\ &\vdots \\ P_{m-1} - Q_{i_m} &= 2P_m, \end{aligned}$$

Dove $Q_{i_1}, \ldots, Q_{i_m}$ sono scelti dall'insieme $\{Q_1, \ldots, Q_n\}$ e dove $P_1, \ldots, P_m$ sono elementi di Γ.

L'idea alla base del procedimento è che finché P_i sarà circa uguale a $2P_{i+1}$, l'altezza di P_i sarà circa il quadruplo di quella di P_{i+1}. Quindi la sequenza di punti $P, P_1, P_2, \ldots$ avrebbero altezza decrescente e potremmo ottenere un insieme di punti con altezza limitata dall'alto. Sfruttando poi la proprietà *a)*, concluderemmo la dimostrazione perché tale insieme sarebbe finito.

Dimostriamolo in maniera rigorosa: dalla prima equazione abbiamo che

$$P = Q_{i_1} + 2P_1.$$

Sostituendo la seconda equazione ($P_1 = Q_{i_2} + 2P_2$) nella prima otteniamo

$$P = Q_{i_1} + 2Q_{i_2} + 4P_2.$$

Iterando il procedimento otteniamo l'uguaglianza

$$P = Q_{i_1} + 2Q_{i_2} + 4Q_{i_3} + \ldots + 2^{m-1}Q_{i_m} + 2^m P_m.$$

La condizione appena scritto ci dice che P è un elemento del sottogruppo generato da $Q_{i_1}, \ldots, Q_{i_m}$ e P_m. Esaminiamo la relazione tra l'altezza di P_{j-1} e di P_j punti della sequenza $P, P_1, P_2, \ldots$, mostrando che l'altezza di P_j sarà considerevolmente più piccola. Applichiamo la proprietà *b)* con $P_0 = -Q_i$ e con costante k_i. Otteniamo allora la disuguaglianza

$$h(P - Q_i) \leq 2h(P) + k_i \quad \forall P \in \Gamma.$$

Sia k' il più grande tra tutti i k_i che si ottengono in questo modo. Allora per $1 \leq i \leq n$ vale:

$$h(P - Q_i) \leq 2h(P) + k', \quad \forall P \in \Gamma.$$

Ciò è possibile perché per la proprietà *d)* sappiamo che i Q_i sono in numero finito. Sia ora k la costante della proprietà *c)*. Allora otteniamo

$$\begin{aligned} 4h(P_j) &\leq h(2P_j) + k \\ &= h(P_{j-1} - Q_{i_j}) + k \\ &\leq 2h(P_{j-1}) + k' + k \end{aligned}$$

che riscriviamo come

$$\begin{aligned} h(P_j) &\leq \frac{1}{2}h(P_{j-1}) + \frac{k' + k}{4} \\ &= \frac{3}{4}h(P_{j-1}) - \frac{1}{4}\left(h(P_{j-1}) - (k' + k)\right). \end{aligned}$$

Se poi $h(P_{j-1}) \geq k' + k$, allora

$$h(P_j) \leq \frac{3}{4}h(P_{j-1}).$$

Quindi nella sequenza $P, P_1, P_2, \ldots$ finché per un j vale la condizione $h(P_{j-1}) \geq k' + k$, allora il punto successivo P_j avrà altezza considerevolmente più piccola, infatti $h(P_j) \leq \frac{3}{4}h(P_{j-1})$. Ma se ripetiamo il procedimento partendo da un altezza fissata qualunque, continuando a moltiplicare per $3/4$, la nuova altezza sarà sempre più prossima allo 0. Eventualmente in questo modo quindi possiamo trovare un indice m per il quale $h(P_m) \leq k' + k$. Se scegliamo m in questo modo otteniamo una stima su P_m che non dipende dal punto iniziale P, e quindi l'insieme finito (per la proprietà *a)*) dei punti con altezza stimata dalla stessa quantità fissata, insieme ai Q_i generano Γ. Infatti ogni punto $P \in \Gamma$ può essere scritto nella forma

$$P = a_1Q_1 + a_2Q_2 + \ldots + a_nQ_n + 2^mR$$

per qualche intero $a_1, \ldots, a_n$ e punto $R \in \Gamma$ che soddisfa la disuguaglianza $h(R) \leq k' + k$. L'insieme di generatori sarà quindi

$$\{Q_1, Q_2, \ldots, Q : n\} \cup \{R \in \Gamma : h(R) \leq k' + k\}$$

che, ricordiamo, è un insieme finito per le proprietà *a)* e *d)*, che conclude la dimostrazione. □

Possiamo formalmente enunciare e dimostrare il seguente

Teorema 4.2 (Teorema di Mordell con un punto razionale di ordine 2)
Sia $\mathcal{C}$ una curva ellittica definita dall'equazione

$$\mathcal{C} : y^2 = x^3 + ax^2 + bx,$$

con $a, b \in \mathbb{Z}$. Allora il gruppo dei punti razionali della curva, $\mathcal{C}(\mathbb{Q})$ è un gruppo abeliano finitamente generato.

Dimostrazione.
Le ipotesi del Teorema (4.1) sono soddisfatte con la funzione altezza, come dimostrato nel precedente capitolo con i lemmi (3.11), (3.12), (3.13), (3.14), quindi $\mathcal{C}(\mathbb{Q})$ è un gruppo abeliano finitamente generato. □

Ci concentreremo per il resto del capitolo nello studio di alcuni esempi in cui utilizzeremo gli strumenti forniti dal percorso svolto per dimostrare il Teorema di Mordell, per descrivere il gruppo dei punti razionali di alcune curve ellittiche. Grazie al Teorema di Mordell sappiamo che tale gruppo Γ sarà abeliano e finitamente generato, e sarà quindi isomorfo alla somma diretta di gruppi ciclici infiniti e di gruppi ciclici finiti di ordine potenza di un primo. Denotiamo con $\mathbb{Z}$ il gruppo additivo degli interi, e con $\mathbb{Z}_m$ il gruppo ciclico $\mathbb{Z}/m\mathbb{Z}$ degli interi modulo m. Allora

$$\Gamma \cong \underbrace{\mathbb{Z} \oplus \mathbb{Z} \oplus \ldots \oplus \mathbb{Z}}_{r \text{ volte}} \oplus \mathbb{Z}_{p_1^{\nu_1}} \oplus \mathbb{Z}_{p_2^{\nu_2}} \oplus \ldots \oplus \mathbb{Z}_{p_s^{\nu_s}}.$$

Equivalentemente, ci dice che esistono dei generatori $P_1, \ldots, P_r, Q_1, \ldots, Q_s \in \Gamma$ tali che ogni punto $P \in \Gamma$ può essere scritto nella forma

$$P = n_1 P_1 + \ldots + n_r P_r + m_1 Q_1 + \ldots + m_s Q_s,$$

dove gli interi n_i, sono unicamente determinati da P, mentre gli interi m_j sono univocamente determinati modulo $p_j^{\nu_j}$.

Definizione 4.1 Sia Γ gruppo dei punti razionali della cubica $\mathcal{C}$ tale che

$$\Gamma \cong \underbrace{\mathbb{Z} \oplus \mathbb{Z} \oplus \ldots \oplus \mathbb{Z}}_{r \text{ volte}} \oplus \mathbb{Z}_{p_1^{\nu_1}} \oplus \mathbb{Z}_{p_2^{\nu_2}} \oplus \ldots \oplus \mathbb{Z}_{p_s^{\nu_s}}.$$

Allora r è detto *rango* di Γ.

Definizione 4.2 Il sottogruppo

$$\mathbb{Z}_{p_1^{\nu_1}} \oplus \mathbb{Z}_{p_2^{\nu_2}} \oplus \ldots \oplus \mathbb{Z}_{p_s^{\nu_s}}$$

di Γ è detto *sottogruppo di torsione* di Γ.

Proviamo adesso a determinare il sottogruppo quoziente $\Gamma/2\Gamma$ di Γ: prima di tutto possiamo scrivere il sottogruppo 2Γ come

$$2\Gamma \cong 2\mathbb{Z} \oplus 2\mathbb{Z} \oplus \ldots \oplus 2\mathbb{Z} \oplus 2\mathbb{Z}_{p_1^{\nu_1}} \oplus 2\mathbb{Z}_{p_2^{\nu_2}} \oplus \ldots \oplus 2\mathbb{Z}_{p_s^{\nu_s}}$$

e quindi il sottogruppo quoziente è nella forma

$$\Gamma/2\Gamma \cong \mathbb{Z}/2\mathbb{Z} \oplus \ldots \oplus \mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}_{p_1^{\nu_1}}/2\mathbb{Z}_{p_1^{\nu_1}} \oplus \ldots \oplus \mathbb{Z}_{p_s^{\nu_s}}/2\mathbb{Z}_{p_s^{\nu_s}}.$$

Ora $\mathbb{Z}/2\mathbb{Z} \cong \mathbb{Z}_2$ è ciclico di ordine 2, invece

$$\mathbb{Z}_{p_i^{\nu_i}}/2\mathbb{Z}_{p_i^{n_i}} \cong \begin{cases} \mathbb{Z}_2 & \text{se } p_i = 2, \\ 0 & \text{se } p_i \neq 2. \end{cases}$$

Quindi possiamo facilmente trovare l'indice di 2Γ in Γ:

$$(\Gamma : 2\Gamma) = 2^{r+(\#\{j:\ p_j=2\})} \tag{4.1}$$

dove con # di un insieme intendiamo la cardinalità di tale insieme, ovvero il numero dei suoi elementi. Denotiamo adesso con $\Gamma[2]$ il sottogruppo formato dai punti $Q \in \Gamma$ tali che $2Q = \mathcal{O}$. Tali punti soddisfano la relazione

$$2\left(n_1P_1 + \ldots + n_rP_r + m_1Q_1 + \ldots + m_sQ_s\right) = \mathcal{O}$$

che accade se $n_i = 0$ per ogni i e se $2m_j \equiv 0 \pmod{p_j^{\nu_j}}$ per ogni j. Se p è dispari e $2m \equiv 0 \pmod{p^\nu}$, allora $m \equiv 0 \pmod{p^\nu}$, mentre se $p = 2$ e $2m \equiv 0 \pmod{p^\nu}$ possiamo concludere che $m \equiv 0 \pmod{p^{\nu-1}}$. Quindi l'ordine del sottogruppo $\Gamma[2]$ sarà

$$\#\Gamma[2] = 2^{\#\{j:\ p_j=2\}}.$$

quindi possiamo riscrivere la relazione (4.2) come

$$(\Gamma : 2\Gamma) = 2^r \cdot \#\Gamma[2]. \tag{4.2}$$

Cerchiamo allora quali sono i punti $Q \in \Gamma$ tali che $2Q = \mathcal{O}$, in modo da vedere che valori può assumere $\#\Gamma[2]$. Tali punti hanno coordinata y nulla come già visto più volte nel precedente capitolo, per cui consideriamo l'equazione della curva

$$\mathcal{C} : y^2 = x^3 + ax^2 + bx.$$

I punti di Γ nella forma $(x_i, 0)$ li otteniamo trovando x_i soluzione di $f(x) = x(x^2+ax+b)$, per cui la prima soluzione sarà per $x_1 = 0$. Per le altre soluzioni, saranno razionali se e solo se il discriminante dell'equazione $x^2 + ax + b$ (ovvero $a^2 - 4b$) sarà un quadrato perfetto, così da ottenere $x_2, x_3 \in \mathbb{Q}$ e quindi $(x_2, 0), (x_3, 0) \in \Gamma$. Osserviamo poi che $\mathcal{O} \in \Gamma[2]$, per cui

$$\#\Gamma[2] = \begin{cases} 2, & \text{se } a^2 - 4b \text{ non è un quadrato perfetto,} \\ 4, & \text{se } a^2 - 4b \text{ è un quadrato perfetto} \end{cases}$$

Ci manca infine da trovare una formula che ci permetta di calcolare il rango di Γ in alcuni casi. Sfruttiamo le mappe $\phi : \Gamma \to \overline{\Gamma}$ e $\psi : \overline{\Gamma} \to \Gamma$ tali che la loro composizione sia la mappa di moltiplicazione per 2 dei punti. Quindi

$$(\Gamma : 2\Gamma) = (\Gamma : \psi \circ \phi(\Gamma)).$$

Sfruttiamo l'inclusione di sottogruppi $\Gamma \supseteq \psi(\overline{\Gamma}) \supseteq 2\Gamma$ per riscrivere

$$(\Gamma : 2\Gamma) = \left(\Gamma : \psi(\overline{\Gamma})\right)\left(\psi(\overline{\Gamma}) : \psi \circ \phi(\Gamma)\right).$$

Studiamo il secondo indice, $\left(\psi(\overline{\Gamma}) : \psi \circ \phi(\Gamma)\right)$ a partire dalla seguente osservazione.

Osservazione 4.1. Siano A un gruppo abeliano, B un suo sottogruppo di indice finito in A, e sia $\psi : A \to A'$ un omomorfismo da A in un qualche gruppo A'. Abbiamo una catena di isomorfismi

$$\frac{\psi(A)}{\psi(B)} \cong \frac{A}{(B + ker(\psi))/B} \cong \frac{A/B}{ker(\psi)/(ker(\psi) \cap B)}.$$

Allora

$$(\psi(A) : \psi(B)) = \frac{(A : B)}{(ker(\psi) : ker(\psi) \cap B)}.$$

Consideriamo ora il caso specifico dell'Osservazione (4.1) con $A = \overline{\Gamma}$ e $B = \phi(\Gamma)$. Allora

$$(\Gamma : 2\Gamma) = \big(\Gamma : \psi(\overline{\Gamma})\big)\big(\psi(\overline{\Gamma}) : \psi \circ \phi(\Gamma)\big) = \big(\Gamma : \psi(\overline{\Gamma})\big) \cdot \frac{\big(\overline{\Gamma} : \phi(\Gamma)\big)}{(ker(\psi) : ker(\psi) \cap \phi(\Gamma))}.$$

Ricordiamo che $\overline{T} \in \phi(\Gamma)$ se e solo se $\overline{b} = a^2 - 4b$ è un quadrato perfetto (Osservazione (3.6)), per cui

$$(ker(\psi) : ker(\psi) \cap \phi(\Gamma)) = \begin{cases} 2, & \text{se } \overline{b} \text{ non è un quadrato perfetto} \\ 1, & \text{se } \overline{b} \text{ è un quadrato perfetto} \end{cases}.$$

Adesso possiamo fare delle buone semplificazioni che ci permettono di non fare distinzione tra il caso in cui $b^2 - 4a$ è un quadrato e il caso in cui non lo è, infatti:

$$2^r = \frac{(\Gamma : 2\Gamma)}{\#\Gamma\,[2]} = \frac{\big(\Gamma : \psi(\overline{\Gamma})\big) \cdot \big(\overline{\Gamma} : \phi(\Gamma)\big)}{4}.$$

Ora calcoliamo gli indici che troviamo al numeratore: troviamo un omomorfismo α tale che

$$\alpha : \Gamma \to \mathbb{Q}^*/\mathbb{Q}^{*2} \quad \text{definito da} \quad \begin{cases} \alpha(x, y) = x \pmod{\mathbb{Q}^{*2}} \\ \alpha(T) = b \pmod{\mathbb{Q}^{*2}} \end{cases}.$$

Abbiamo mostrato nella Proposizione (3.17) che il nucleo di α è l'immagine di $\psi(\overline{T})$, ovvero è isomorfa a

$$\alpha(\Gamma) \cong \Gamma/ker(\alpha) \cong \Gamma/\psi(\overline{\Gamma}).$$

Dunque $\big(\Gamma : \psi(\overline{\Gamma})\big) = \#\alpha(\Gamma)$. Similmente poi troviamo, a partire dall'omomorfismo $\overline{\alpha} : \overline{\Gamma} \to \mathbb{Q}^*/\mathbb{Q}^{*2}$, che $\big(\overline{\Gamma} : \phi(\Gamma)\big) = \#\overline{\alpha}(\overline{\Gamma})$. Sostituendo i risultati appena ottenuti nella formula precedente otteniamo una formula alternativa per il rango di Γ:

$$2^r = \frac{\#\alpha(\Gamma) \cdot \#\overline{\alpha}(\overline{\Gamma})}{4}.$$

Ci rimane quindi da trovare un metodo per determinare l'immagine di α, e quindi quali punti razionali modulo $\mathbb{Q}^{*2}$ possono essere coordinate razionali di punti di Γ. Scriviamo allora

$$x = \frac{m}{e^2}, \quad y = \frac{n}{e^3}$$

ridotti ai minimi termini e con $e > 0$.

Se $m = 0$, allora $(x, y) = T$ e $\alpha(T) = b$. Dunque $b \pmod{\mathbb{Q}^{*2}}$ è sempre un elemento di $\alpha(\Gamma)$. Se $a^2 - 4b$ è un quadrato perfetto, diciamo $a^2 - 4b = d^2$, allora in Γ ci sono altri due punti di ordine 2, ovvero

$$\left(\frac{-a+d}{2}, 0\right) \quad \text{e} \quad \left(\frac{-a-d}{2}, 0\right).$$

Quindi se $a^2 - 4b = d^2$, $\alpha(\Gamma)$ contiene $(-a \pm d)/2$.

Ora studiamo i punti con $m, n \neq 0$. Per tali punti vale

$$n^2 = m^3 + am^2e^2 + bme^4 = m(m^2 + ame^2 + be^4).$$

Con gli stessi argomenti che abbiamo utilizzato nella dimostrazione della Proposizione (3.17), se $MCD(m, b) = b_1$, allora possiamo scrivere che

$$m = m_1 b_1 \quad \text{e} \quad b = b_1 b_2$$

con $MCD(m_1, b_2) = 1$ e $m_1 > 0$. Sostituendo i valori nell'equazione della curva otteniamo

$$n^2 = b_1 m_1 (b_1^2 m_1^2 + a b_1 m_1 e^2 + b_1 b_2 e^4) = b_1^2 m_1 (b_1 m_1^2 + a m_1 e^2 + b_2 e^4).$$

Quindi $b_1^2 | n^2$, ovvero $b_1 | n$. Riscrivendo $n = b_1 n_1$ e dividendo l'equazione intera per b_1^2 otteniamo

$$n_1^2 = m_1 (b_1 m_1^2 + a m_1 e^2 + b_2 e^4)$$

e nuovamente le quantità m_1 e $b_1 m_1^2 + a m_1 e^2 + b_2 e^4$ sono prime tra loro. Ma allora entrambe devono essere dei quadrati, essendo il loro prodotto un quadrato perfetto, quindi possiamo scrivere $n_1 = NM$ con M, N tali che

$$M^2 = m_1, \qquad N^2 = b_1 m_1^2 + a m_1 e^2 + b_2 e^4,$$

da cui

$$N^2 = b_1 M^4 + a M^2 e^2 + b_2 e^4.$$

Questo significa che se un punto $(x, y) \in \Gamma$ è tale che $y \neq 0$, allora lo si può riscrivere nella forma

$$x = \frac{b_1 M^2}{e^2}, \qquad y = \frac{b_1 MN}{e^3},$$

per cui le ascisse di ogni punto della curva, modulo $\mathbb{Q}^{*2}$ sono i valori possibili di b_1, ed essendo b_1 un divisore dell'intero $b \neq 0$, ci sono solo un numero finito di possibilità per b_1.

Vediamo quindi come calcolare l'ordine di $\alpha(\Gamma)$: Scegliamo un intero b e fattorizziamolo come prodotto $b = b_1 b_2$ in tutti i modi possibili. Per ogni modo possibile, scriviamo l'equazione nelle variabili M, e, N fissati a, b_1, b_2

$$N^2 = b_1 M^4 + a M^2 e^2 + b_2 e^4. \tag{4.3}$$

Allora $\alpha(\Gamma)$ consiste in $b \pmod{\mathbb{Q}^{*2}}$ insieme ai $b_1 \pmod{\mathbb{Q}^{*2}}$ per cui esiste una soluzione con M non nullo.

Notiamo che il procedimento tiene conto anche dei valori nel caso in cui $a^2 - 4b = d^2$, infatti in quel caso possiamo fattorizzare b come

$$n = \frac{-a+d}{2} \cdot \frac{-a-d}{2},$$

quindi col procedimento appena svolto andremmo a considerare l'equazione

$$N^2 = \left(\frac{-a \pm d}{2}\right) M^4 + aM^2e^2 + \left(\frac{-a \mp d}{2}\right) e^4,$$

che ha soluzione banale $(M, e, N) = (1, 1, 0)$.

Quindi, per determinare l'ordine di $\alpha(\Gamma)$ consideriamo tutte le possibili fattorizzazioni di b come prodotto di due interi, verifichiamo se l'equazione che ne ricaviamo ha soluzioni con $M \neq 0$ e ogni volta che si verificano queste condizioni otteniamo un nuovo punto nella forma

$$x = \frac{b_1 M^2}{e^2}, \quad y = \frac{b_1 MN}{e^3}.$$

Il motivo per cui questo metodo può dare dei problemi nella risoluzione è che non esiste un metodo generale per determinare se l'equazione nella forma (4.3) abbia o meno soluzione.

Passiamo ora all'applicazione del metodo appena illustrato in una serie di esempi, in cui calcoleremo il rango di Γ. Notiamo che per determinare totalmente la struttura del gruppo $\mathcal{C}(\mathbb{Q})$ avremmo bisogno anche di determinare la struttura del sottogruppo di torsione, che richiede degli argomenti riguardo ai punti di ordine finito della cubica come il Teorema di Nagell-Lutz, che noi però non tratteremo. Ci concentreremo in particolare infatti solo sullo studio della curva con gli strumenti forniti dalla dimostrazione del Teorema di Mordell, per mettere in luce la loro importanza nella risoluzione effettiva di queste equazioni.

Esempio 4.1 Sia $\mathcal{C}$ la curva di equazione

$$\mathcal{C} : y^2 = x^3 - 5x.$$

Essendo che $a = 0$, $b = -5$, abbiamo solamente 4 possibilità per b_1, ovvero $b_1 = 1, -1, 5, -5$. Le equazioni corrispondenti a tali valori sono

$$\begin{aligned} (i) \quad & N^2 = M^4 - 5e^4 \\ (ii) \quad & N^2 = -M^4 + 5e^4 \\ (iii) \quad & N^2 = 5M^4 - e^4 \\ (iv) \quad & N^2 = -5M^4 + e^4. \end{aligned}$$

Notiamo che le equazioni (i) e (ii) sono uguali rispettivamente alle equazioni (iii) e (iv), ma con le variabili e ed M invertite. Nel nostro caso troveremo soluzioni con entrambi e ed M non nulli, per cui possiamo dire che una ammette soluzione accettabile se e solo se la ammette l'altra. Troviamo allora le soluzioni $(M, N, e) = (3, 1, 2)$ per (i) e $(1, 2, 1)$ per (ii), per cui troviamo i punti razionali cercati a partire dalle relazioni

$$x = \frac{b_1 M^2}{e^2}, \quad y = \frac{b_1 MN}{e^3},$$

trovando i punti $\left(\frac{9}{4}, \frac{3}{8}\right)$ e $(-1, -2)$. Quindi

$$\alpha(\Gamma) = \{\pm 1, \pm 5\} \pmod{\mathbb{Q}^{*2}}.$$

Ora è necessario cercare allo stesso modo le soluzioni per

$$\overline{\mathcal{C}} : y^2 = x^3 + 20x$$

per trovare $\overline{\alpha}(\overline{\Gamma})$. Essendo $\overline{b} = a^2 - 4b = 20$, le possibili scelte per $\overline{b_1}$ sono

$$\overline{b_1} = \pm 1, \pm 2, \pm 4, \pm 5, \pm 10, \pm 20.$$

Osserviamo che dalla relazione $\overline{b_1} \cdot \overline{b_2} = \overline{b} = 20$ deduciamo che $\overline{b_1}$ e $\overline{b_2}$ hanno stesso segno. Se sono entrambi negativi, allora l'equazione

$$N^2 = \overline{b_1} M^4 + \overline{b_2} e^4$$

non ha soluzioni reali non nulle, per cui non ne ha neanche di razionali. Segue che

$$\overline{\alpha}(\overline{\Gamma}) \subseteq \{1, 2, 4, 5, 10, 20\}.$$

Dopodiché, notiamo che

$$\begin{aligned} \overline{\alpha}(\overline{\mathcal{O}}) &= 1 \equiv 4 \pmod{\mathbb{Q}^{*2}}, \\ \overline{\alpha}(\overline{T}) &= \overline{b} = 20 \equiv 5 \pmod{\mathbb{Q}^{*2}}, \end{aligned}$$

quindi sono elementi di $\overline{\alpha}(\overline{\Gamma})$. Rimangono da verificare solo i valori di $\overline{b_1} = 2, 10$. Cerchiamo allora soluzioni dell'equazione

$$N^2 = 2M^4 + 10e^4.$$

Ricordiamo che M, N, e sono coprimi, ma se M fosse multiplo di 5 lo dovrebbe essere anche N, che contraddice la coprimità. Quindi M è coprimo con 5 e per il piccolo teorema di Fermat vale che $M^4 \equiv 1 \pmod 5$. Riducendo l'intera equazione modulo 5 vediamo che necessariamente $N^2 \equiv 2 \pmod 5$, che non ha soluzioni, da cui ricaviamo che $2 \notin \overline{\alpha}(\overline{\Gamma})$. Ora, essendo $5 \in \overline{\alpha}(\overline{\Gamma})$ e $2 \notin \overline{\alpha}(\overline{\Gamma})$, dovendo essere quest'ultimo un gruppo, sicuramente $10 \notin \overline{\alpha}(\overline{\Gamma})$, altrimenti sarebbe possibile ottenere 2 come prodotto tra 10 e l'inverso di 5. Concludiamo quindi dicendo che

$$\overline{\alpha}(\overline{\Gamma}) = \{1, 5\} \pmod{\mathbb{Q}^{*2}}$$

e quindi

$$2^r = \frac{\#\alpha(\Gamma) \cdot \#\overline{\alpha}(\overline{\Gamma})}{4} = \frac{4 \cdot 2}{4} = 2,$$

per cui il rango di Γ è 1.

Vedendo l'esempio precedente, è chiaro che gli strumenti che abbiamo trovato siano piuttosto raffinati. Ne esistono però alcune per cui non si riescono a scartare alcuni fattori possibili b_1 tramite l'utilizzo dei moduli e per cui non si riesce a trovare neanche soluzione per l'equazione associata. Questi casi rappresentano esattamente il problema che avevamo posto in precedenza, che è il motivo per cui in generale non è sempre possibile trovare l'insieme di generatori del gruppo. In generale è possibile costruire facilmente degli esempi in cui il rango del gruppo Γ sia $0, 1, 2$ (ad esempio per la cubica $y^2 = x^3 + px$ con p un numero primo) o altri valori piccoli, ma diventa molto più complesso costruire esempi di rango arbitrariamente grande, infatti rimane tutt'oggi un problema aperto la possibilità di trovare curve con rango arbitrariamente grande. Inoltre, anche curve con coefficienti piccoli possono avere come generatori punti di altezza molto grande. Un esempio è la curva $y^2 = x^3 + 877x$, il cui gruppo Γ è generato da $T = (0, 0)$ e $P = (x_0, y_0)$ dove

$$x_0 = \left(\frac{612776083187947368101}{78841535860683900210}\right)^2.$$

Calcoliamo ora il rango per altre curve, col metodo sviluppato.

Esempio 4.2 Siano date le curve

$$\mathcal{C}_p : y^2 = x^3 + px, \quad \overline{\mathcal{C}}_p : y^2 = x^3 - 4px$$

In generale i possibili valori per b_1 saranno $\pm 1, \pm p$, mentre quelli di $\overline{b_1}$ saranno ± 1, ± 2, ± 4, $\pm p$, $\pm 2p$, $\pm 4p$. Le corrispondenti equazioni per i valori di b_1 sono

$$\begin{aligned}
&(i) \quad N^2 = M^4 + pe^4 \\
&(ii) \quad N^2 = pM^4 + e^4 \\
&(iii) \quad N^2 = -M^4 - pe^4 \\
&(iv) \quad N^2 = -pM^4 - e^4.
\end{aligned}$$

Vediamo subito che le equazioni (iii) e (iv) non possono ammettere soluzioni razionali, perché non ne ammettono di reali. Invece, per quanto riguarda 1 e p sappiamo che

$$\begin{aligned}
\alpha(\mathcal{O}) &= 1 \pmod{\mathbb{Q}^{*2}}, \\
\alpha(T) &= b = p \pmod{\mathbb{Q}^{*2}},
\end{aligned}$$

da cui $\alpha(\Gamma) = \{1, p\}$ indipendentemente da p. Inoltre per $\overline{\alpha}$ possiamo osservare similmente che

$$\begin{aligned}
\overline{\alpha}(\overline{\mathcal{O}}) &= 1 \equiv 4 \pmod{\mathbb{Q}^{*2}}, \\
\overline{\alpha}(\overline{T}) &= \overline{b} = -4p \equiv -p \pmod{\mathbb{Q}^{*2}},
\end{aligned}$$

per cui vale $\{1, -p\} \subseteq \overline{\alpha}(\overline{\Gamma})$. Notiamo che se $-1, +2 \in \overline{\alpha}(\overline{\Gamma})$, allora anche

$$\begin{aligned}
-2 &= (-1) \cdot 2, \\
+p &= (-1) \cdot (-p), \\
\pm 2p &= (\pm 2) \cdot p, \\
\pm 4p &\equiv \pm p \pmod{\mathbb{Q}^{*2}}, \\
\pm 4 &\equiv \pm 1 \pmod{\mathbb{Q}^{*2}},
\end{aligned}$$

sono elementi dell'immagine, essendo $\overline{\alpha}$ un omomorfismo. Allora potremmo dire che $\overline{\alpha}(\overline{\Gamma}) = \{\pm 1, \pm 2, \pm p, \pm 2p\}$. In questo caso avremmo

$$2^r = \frac{\#\alpha(\Gamma) \cdot \#\overline{\alpha}(\overline{\Gamma})}{4} = \frac{2 \cdot 8}{4} = 4$$

e Γ avrebbe rango 2, che sarà il massimo rango ammissibile per curve di questo tipo di conseguenza. Similmente possiamo mostrare che il rango di Γ può anche essere 0 o 1 in base alla scelta di p. Facciamo un esempio per ognuno di questi tre ranghi.

i) **Rango = 0**: $p = 7$.
Se $b_1 \in \{-1, -2, -4\}$, otteniamo l'equazione

$$N^2 = b_1 M^4 - \frac{28}{b_1} e^4$$

che ridotta modulo 7 risulta essere

$$N^2 \equiv b_1 M^4 \pmod 7.$$

Osserviamo che i resti possibili sia per i quadrati, sia per le potenze quarte modulo 7 sono $0, 1, 2, 4$. Se uno tra M, N fosse divisibile per 7 dovrebbe esserlo anche l'altro, il che contraddice la coprimalità dei due. Quindi entrambi possono assumere solamente i valori $1, 2, 4$. Notiamo poi che se $M^4 \in \{1, 2, 4\} \subseteq \mathbb{Z}_7$, allora $b_1 M^4 \in \{3, 5, 6\} \subseteq \mathbb{Z}_7$, quindi l'equivalenza non può avere soluzioni. Segue che $-1, -2 \notin \overline{\alpha}(\overline{\Gamma})$. Dunque anche $p, 2p \notin \overline{\alpha}(\overline{\Gamma})$, altrimenti $\overline{\alpha}$ non sarebbe un omomorfismo.

Ci è sufficiente adesso mostrare che $2 \notin \overline{\alpha}(\overline{\Gamma})$, sempre per il fatto che $\overline{\alpha}$ deve essere un omomorfismo, e quindi $-2p = 2 \cdot (-p)$ è un elemento dell'immagine se e solo se lo è 2. L'equazione che otteniamo è

$$N^2 = 2M^4 - 14e^4.$$

Vediamo subito che N è pari, e quindi M, e devono essere dispari per la condizione di coprimalità. Sia $n \in \mathbb{Z}$ tale che $N = 2n$. Sostituendo nell'equazione ricaviamo

$$\begin{aligned} 4n^2 &= 2M^4 - 14e^4 \\ 2n^2 &= M^4 - 7e^4 \\ 2n^2 - 2e^4 &= M^4 - 9e^4 \\ 2(n - e^2)(n + e^2) &= (M^2 - 3e^2)(M^2 + 3e^2). \end{aligned}$$

Essendo M, e dispari, i termini $M^2 \pm 3e^2$ sono entrambi pari, di conseguenza lo devono essere anche i termini $n \pm e^2$, altrimenti a sinistra dell'uguaglianza avrei un numero non divisibile per 4, mentre a destra avrei un multiplo di 4. Quindi n deve essere dispari. Notiamo che tra $n + e^2$ e $n - e^2$, uno deve essere multiplo di 2 ma non di 4, mentre l'altro deve essere necessariamente multiplo di 4. Se così non fosse, i due termini dovrebbero necessariamente essere congrui tra loro modulo 4 (o sono entrambi congrui a 2, o entrambi a 0). Ma se $n - e^2 \equiv n + e^2 \pmod 4$, allora $2e^2 \equiv 0 \pmod 4$, per cui sarebbe necessario che e sia pari, che non è. l'analogo ragionamento può essere applicato ai termini $M^2 \pm 3e^2$. In particolare qua possiamo essere più precisi: $M^2 \equiv e^2 \equiv 1 \pmod 4$, perché entrambi

sono dispari. Per cui $M^2 - 3e^2 \equiv 2 \pmod 4$. Per quanto concerne l'altro termine invece, osserviamo che anche modulo 8 l'unico residuo quadratico dispari è 1. Ma allora

$$M^2 + 3e^4 \equiv 4 \pmod 8.$$

Ciò significa che il prodotto $(M^2 + 3e^2)(M^2 - 3e^2)$ ha esattamente tre fattori 2, mentre $2(n - e^2)(n + e^2)$ ne ha sicuramente almeno quattro. Ciò è sufficiente a concludere che l'equazione non ammette soluzioni ammissibili, per cui le immagini di $\alpha(\Gamma)$ e $\overline{\alpha}(\overline{\Gamma})$ sono le più piccole possibili, e il rango dato dalla formula

$$2^r = \frac{2 \cdot 2}{4} = 1$$

è esattamente 0.

ii) **Rango = 1**: $p = 3$.
Mostriamo che $-1, 2 \notin \overline{\alpha}(\overline{\Gamma})$. Se $b_1 \in \{-1, 2, -4 \equiv -1\}$, allora riduciamo la solita equazione con $p = 3$ modulo 3, ottenendo

$$N^2 \equiv 2M^4 \pmod 3$$

che ha soluzione solo per $M, N \equiv 0 \pmod 3$ e che quindi non sarebbero coprimi. Invece vediamo facilmente che $-2 \in \overline{\alpha}(\overline{\Gamma})$, infatti la terna $(M, N, e) = (1, 2, 1)$ soddisfa l'equazione corrispondente

$$N^2 = 4 = (-2) \cdot 1^4 + 6 \cdot 1^4 = -2M^4 + 6e^4.$$

Allora sfruttando la proprietà di omomorfismo di $\overline{\alpha}$ vediamo che

$$\overline{\alpha}(\overline{\Gamma}) = \{1, -p, -2, 2p = (-2) \cdot (-p)\},$$

da cui concludiamo

$$2^r = \frac{2 \cdot 4}{4} = 2$$

e quindi $r = 1$.

iii) **Rango = 2**: $p = 73$.
In questo caso specifico abbiamo la possibilità di affrontare una questione che è utile vedere per fare chiarezza: come detto all'inizio, il rango di $\mathcal{C}_{73}$ sarà 2 se e solo se l'immagine di $\overline{\alpha}(\overline{\Gamma})$ contiene tutti e 8 gli elementi. Però se proviamo a calcolare l'equazione corrispondente per -1, ovvero

$$N^2 = -M^4 + 292e^4$$

notiamo subito che riducendola modulo 4 otteniamo $N^2 + M^4 \equiv 0 \pmod 4$, che ha soluzione solo per $N^2, M^4 \equiv 0 \pmod 4$ e quindi sia N sia M dovranno essere pari contraddicendo la coprimalità dei due. Sembra allora che $-1 \notin \overline{\alpha}(\overline{\Gamma})$, e quindi il rango della curva non potrà essere 2. In realtà ricordiamo che stiamo considerando solo gli 8 valori

$\pm 1, \pm 2, \pm p, \pm 2p$ come possibili elementi dell'immagine perché gli elementi di $\overline{\alpha}(\overline{\Gamma})$ li consideriamo modulo $\mathbb{Q}^{*2}$, ma stiamo tralasciando la verifica di alcuni casi possibili per b_1 divisore di b. L'equazione per -4, ovvero

$$N^2 = -4M^4 + 73e^4$$

ha infatti soluzione $(M, N, e) = (2, 3, 1)$ accettabile, quindi $-4 \in \overline{\alpha}(\overline{\Gamma})$ e così anche $-1 \in \overline{\alpha}(\overline{\Gamma})$. Nei punti precedenti dell'esempio abbiamo infatti verificato per tutti i valori possibili che non appartenessero all'immagine di $\overline{\alpha}$ perché non è in generale vero che se l'equazione per un valore non ha soluzioni ammissibili, non le ha anche per tutti i valori della sua stessa classe di resto modulo $\mathbb{Q}^{*2}$. Il controesempio è il caso attuale $p = 73$.

Concludiamo osservando che l'equazione per -2:

$$N^2 = -2M^4 + 146e^4$$

ha soluzione accettabile $(M, N, e) = (1, 1, 12)$, per cui $-2 \in \overline{\alpha}(\overline{\Gamma})$. Sfruttando le proprietà di omomorfismo di $\overline{\alpha}$ quindi si verifica facilmente che $\overline{\alpha}(\overline{\Gamma}) = \{\pm 1, \pm 2, \pm p, \pm 2p\}$, da cui otteniamo che il rango di $\mathcal{C}_{73}$ è esattamente 2.

Riguardo alle curve $\mathcal{C}_p$ di questa forma ci sono dei risultati che ci permettono di ottenere da subito informazioni precise sul rango, mentre in altri casi ci sono solo congetture. Ad esempio si può mostrare che se $p \equiv 7$ *o* $11 \pmod{16}$, il rango della curva è sempre 0. Invece è solamente una congettura al momento che se $p \equiv 1 \pmod 8$ il rango è solamente 0 o 2, ed esistono casi per entrambi i ranghi. Abbiamo visto prima che $\mathcal{C}_{73}$ ha rango 2, mentre è possibile mostrare che $\mathcal{C}_{17}$ e $\mathcal{C}_{43}$ ad esempio hanno rango nullo. Infine rimangono delle congetture anche per le curve di rango 1. Si pensa infatti che se $p \equiv 3, 5, 13$ *o* $15 \pmod{16}$ il rango sia esattamente 1.

Vediamo adesso un caso molto simile al precedente in cui riusciamo a dare delle informazioni in generale su una classe di curve.

Esempio 4.3 Sia $p \neq 2$ un numero primo. Sia data la curva $\mathcal{D}_p$ di equazione

$$\mathcal{D}_p : y^2 = x^3 - px$$

e la corrispondente curva

$$\overline{\mathcal{D}}_p : y^2 = x^3 + 4px$$

In questo caso possiamo ottenere delle informazioni molto importanti a prescindere dalla scelta di p. Capiamo prima di tutto che insieme è $\alpha(\Gamma) \subseteq \{\pm 1, \pm p\}$. Come al solito, conosciamo l'immagine dei punti $\mathcal{O}$ e T, quindi

$$\alpha(\mathcal{O}) = 1$$
$$\alpha(T) = -p$$

per cui

$$\{1, -p\} \subseteq \alpha(\Gamma) \subseteq \{\pm 1, \pm p\}.$$

Ora guardiamo le equazioni che ci vengono fornite dai valori $b_1 = -1, p$. Chiaramente $-1 \in \alpha(\Gamma)$ se e solo se $p \in \alpha(\Gamma)$ dovendo essere α un omomorfismo, inoltre le due equazioni

sono identiche a meno di scambiare M ed e, e una soluzione qualunque accettabile per entrambe le equazioni deve soddisfare la relazione $Me \neq 0$, quindi consideriamo solo quella per -1. L'equazione è

$$N^2 = -M^4 + pe^4.$$

Notiamo subito che se $p \equiv 3 \pmod 4$, allora guardando l'equazione modulo 4 otteniamo che $n^2 \equiv -M^4 \pmod 4$, che ha soluzione solo per M, N pari. Quindi per $p \equiv 3 \pmod 4$, abbiamo mostrato che $\alpha(\Gamma) = \{1, -p\}$.

Un altro modo per avere la stessa informazione poteva sfruttare invece le relazioni tra residui quadratici e moduli, e in un qualche modo quindi possiamo essere aiutati dalla legge di reciprocità quadratica di Gauss, che in questo caso ci darà un'informazione già ricavata ma più avanti potrà essere d'aiuto per fare una considerazione ulteriore. Utilizzando per notazione il simbolo di Legendre, ricordiamo che

$$\left(\frac{-1}{p}\right) = \begin{cases} 1, & \text{se } p \equiv 1 \pmod 4 \\ -1, & \text{se } p \equiv 3 \pmod 4 \end{cases}.$$

Ora riprendiamo la nostra equazione

$$N^2 = -M^4 + pe^4.$$

Necessariamente p non deve dividere nè N nè M, altrimenti dividerebbe entrambi e non sarebbero più coprimi. Ma allora riducendo l'equazione modulo p vediamo che

$$\begin{aligned} N^2 &\equiv -M^4 \pmod p \\ \left(\frac{N}{M^2}\right)^2 &\equiv -1 \pmod p \end{aligned}$$

e quindi -1 è un residuo quadratico modulo p, da cui l'informazione di prima. Inoltre, possiamo ulteriormente procedere mostrando che se $p \equiv 13 \pmod{16}$, allora l'equazione

$$N^2 = -M^4 + pe^4$$

non ammette soluzioni. Infatti i possibili residui quadratici modulo 16 sono $\{0, 1, 4, 9\}$ e le possibili potenze quarte solamente 0 e 1. Se riduciamo l'equazione modulo 16 quindi vediamo subito che non ci sono possibili scelte di tali valori in modo che valga una congruenza modulo 16, quindi non esistono soluzioni intere, da cui $\alpha(\Gamma) = \{1, -p\}$.

Cerchiamo ora di ottenere informazioni su $\overline{\alpha}(\overline{\Gamma}) \subseteq \{\pm 1 \equiv \pm 4, \pm 2, \pm p \equiv \pm 4p, \pm 2p\}$ in cui tutte le equivalenze sono sempre modulo $\mathbb{Q}^{*2}$. Come prima

$$\begin{aligned} &\overline{\alpha}(\overline{\mathcal{O}}) = 1 \\ &\overline{\alpha}(\overline{T}) = 4p \equiv p \pmod{\mathbb{Q}^{*2}}. \end{aligned}$$

Notiamo che poi che per ogni scelta di $b_1 < 0$, sarà anche $b_2 < 0$. Quindi l'equazione

$$N^2 = b_1 M^4 + b_2 e^4$$

non avrà soluzioni accettabili, essendo N^2 non negativo e il termine invece a destra dell'uguaglianza sempre negativo o nullo, e si annulla solo per $M, e = 0$. Quindi tutti gli elementi $-1, -2, -p, -2p$ non saranno elementi di $\overline{\alpha}(\overline{\Gamma})$. Sappiamo per adesso che

$$\{1, p\} \subseteq \overline{\alpha}(\overline{\Gamma}) \subseteq \{1, p, 2, 2p\}\,.$$

Come prima ci basta considerare uno tra $b_1 = 2$ e $b_1 = 2p$ per determinare l'insieme. L'equazione per $b_1 = 2$ è

$$N^2 = 2M^4 + 2pe^4.$$

Sfruttiamo questa volta fino da subito la legge di reciprocità quadratica di Gauss e le sue conseguenze. In particolare sappiamo che

$$\left(\frac{2}{p}\right) = \begin{cases} 1, & \text{se } p \equiv 1, 7 \pmod 8 \\ -1, & \text{se } p \equiv 3, 5 \pmod 8 \end{cases}.$$

Riducendo la nostra equazione modulo p infatti, osservando che M ed N devono essere nuovamente coprimi con p, possiamo scrivere

$$N^2 \equiv 2M^4 \pmod p$$
$$\left(\frac{N}{M^2}\right)^2 \equiv 2 \pmod p$$

per cui 2 è un residuo quadratico modulo p, che dovrà essere necessariamente congruo a 1 o 7 modulo 8 affinché l'equazione possa ammettere delle soluzioni. Si può trovare, come prima lavorando con le congruenze modulo 8 lo stesso fatto, ma la legge di reciprocità quadratica ci dà direttamente l'informazione cercata. Abbiamo già allora un risultato interessante: Se $p \equiv 3 \pmod 8$, o $p \equiv 13 \pmod{16}$, allora

$$\alpha(\Gamma) = \{1, -p\}$$
$$\overline{\alpha}(\overline{\Gamma}) = \{1, p\}$$

per cui il rango di $\mathcal{D}_p$ sarà 0.

Anche in questo caso inoltre sappiamo che il massimo rango di una qualunque curva $\mathcal{D}_p$ è 2, e può essere esattamente 2 se e solo se $p \equiv 1 \pmod 8$. Osserviamo anche che se $p \equiv 5 \pmod 8$ non è in generale possibile stabilire se il suo rango sia 0 o 1 da questa congruenza, infatti nell'Esempio (4.1) il rango è esattamente 1, mentre sappiamo che se p ad esempio è 13, allora il rango sarà 0. Troviamo esplicitamente un esempio in cui il rango sia esattamente 2.

i) **Rango = 2**: $p = 17$.
L'equazione

$$N^2 + M^4 = 17e^4$$

ha soluzione accettabile $(M, N, e) = (2, 1, 1)$ ad esempio, quindi $\alpha(\Gamma) = \{\pm 1, \pm 17\}$. L'equazione

$$N^2 = 2M^4 + 34e^4$$

ha invece soluzione ammissibile $(M, N, e) = (1, 6, 1)$, da cui $\overline{\alpha}(\overline{\Gamma}) = \{\pm 1, \pm 17\}$ e quindi il rango di $\mathcal{D}_{17}$ è esattamente 2.

Una cosa interessante che si può verificare con un programma di calcolo, è che verificando per qualunque numero primo $p \equiv 1 \pmod 8$ fino a 100 si vede che la prima delle due equazioni ammette sempre soluzioni accettabili (e sembra anche che in questi casi M debba essere una potenza di 2), mentre per $p = 113$ per valori di M, N, e minori di 2000 non si trovano soluzioni ammissibili. Ora $17 \equiv 113 \pmod{32}$, il che ci suggerisce che se l'equazione per $p = 113$ non avesse effettivamente soluzioni (come può sembrare realistico credere) stabilire un criterio per individuare se il rango della curva sia 2 o meno richiederebbe almeno di dover sfruttare le congruenze modulo 64, quindi sembrerebbe essere già molto più complesso dei risultati trovati fino ad adesso.

ii) $p \equiv 7 \pmod 8$.

In questo caso abbiamo visto che $\alpha(\Gamma) = \{1, -p\}$, infatti se $p \equiv 7 \pmod 8$ necessariamente $p \equiv 3 \pmod 4$. Per determinare il rango delle curve $\mathcal{D}_p$ quindi dobbiamo trovare $\overline{\alpha}(\overline{\Gamma})$, per cui abbiamo bisogno di studiare l'equazione

$$N^2 = 2M^4 + 2pe^4.$$

Prima di tutto troviamo soluzioni esplicite (M, N, e) in alcuni casi significativi: consideriamo $p = 7, 47, 23, 31$.

- Se $p = 7$: $(1, 4, 1)$ è soluzione ammissibile.
- Se $p = 47$: $(3, 16, 1)$ è soluzione ammissibile.
- Se $p = 23$: $(5, 36, 1)$ è soluzione ammissibile.
- Se $p = 31$: $(1, 8, 1)$ è soluzione ammissibile.

In tutti questi quattro casi quindi avremo che $\overline{\alpha}(\overline{\Gamma}) = \{\pm 1, \pm p\}$ e il rango di conseguenza sarà 1. Il motivo per cui sono casi significativi lo possiamo trovare negli esempi precedenti: per trovare una dimostrazione del fatto che alcune equazioni non avessero soluzioni abbiamo sfruttato principalmente le congruenze modulo p e modulo potenze di 2. I valori di p che abbiamo scelto sono infatti a due a due distinti modulo 32 (e quindi sono tutti i valori possibili per p modulo 32). Per provare ad affermare quindi che per alcune scelte di p l'equazione associata non ammetta soluzione sfruttando le potenze di 2, abbiamo bisogno di valutare l'equazione modulo 64 almeno, il che ci porta a pensare che il criterio utilizzato spesso fino ad adesso sembra non dare informazioni utili questa volta.

L'opzione che sembra più naturale allora è che l'equazione possa sempre ammettere soluzioni. Provare a dimostrare un risultato simile per un $p \equiv 7 \pmod 8$ generico però richiede uno studio più approfondito dell'equazione. Provando a verificare questa tesi con dei programmi di calcolo, otteniamo soluzioni rapidamente per tutti i valori di p che stiamo prendendo in esame, fino a $p = 199$. Infatti per $p = 223$ si trova che non esistono soluzioni ammissibili dell'equazione con $M, N < 2000$. Potrebbero quindi esistere soluzioni per M ed N più grandi, o invece l'equazione potrebbe non ammettere soluzione per questo valore specifico, o più generalmente forse per ogni $p \equiv 223 \equiv 95 \pmod{128}$.

Concludiamo l'esempio allora non trovando una soluzione a questo caso, ma avendone visto alcune possibili.

Osserviamo che in molti di questi esempi abbiamo applicato uno strumento che è risolutivo in alcuni casi, ma che non sempre ci permette di avere l'informazione cercata. Nel momento in cui riducendo l'equazione modulo potenze di 2 e modulo p non si ottengono informazioni cruciali, infatti, la risoluzione diventa molto più complessa.

Abbiamo visto quindi molti esempi di calcolo di rango di queste curve. Abbiamo scelto equazioni della forma $y^2 = x^3 \pm px$ perché sono la classe più semplice di curve ellittiche che possiamo affrontare, avendo il minor numero di casi possibili da studiare e dandoci già queste un'idea molto accurata di come è poi possibile affrontare anche curve di complessità più alta. Concludiamo così la sezione di studio delle curve ellittiche con l'utilizzo del Teorema di Mordell.

Bibliografia

[1] Titu Andreescu, Dorin Andrica, Ion Cucurezeanu et al. *An introduction to Diophantine equations: A problem-based approach*. Springer, 2010.

[2] Maurizio Cailotto. *Curve algebriche piane*. 2013.

[3] John William Scott Cassels. «Diophantine equations with special reference to elliptic curves». In: *Journal of the London Mathematical Society* 1.1 (1966), pp. 193–291.

[4] Joseph H Silverman e John Torrence Tate. *Rational points on elliptic curves*. Vol. 9. Springer, 1992.

[5] Irene Udassi. *Il gruppo dei punti razionali di una curva ellittica*. 2014.