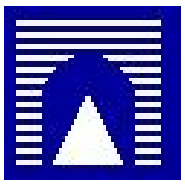
UNIVERSITÀ DEGLI STUDI ROMA TRE
FACOLTÀ DI S.M.F.N.



Tesi di Laurea in Matematica
presentata da
Flaminia Susco

Risoluzione dell'equazione di Pell

Relatore
Prof.ssa Francesca Tartarone

Il Candidato Il Relatore

ANNO ACCADEMICO 2005-2006
Ottobre 2006

Indice

Introduzione

Si dice equazione di Pell un'equazione diofantea del tipo:

$$x^2 - dy^2 = 1 \qquad (\ o,\ \textit{equivalentemente}\ \ x^2 = dy^2 + 1) \tag{0.0.1}$$

dove d è un numero naturale che non è un quadrato perfetto.[1]
È un fatto notevole che l'equazione di Pell ammette sempre una soluzione (x, y) nell'insieme dei numeri naturali, e che tali soluzioni siano infinite.
Riferimenti a singoli casi di equazioni di Pell si trovano sparsi in tutta la storia della matematica. La più curiosa di queste citazioni avviene nel cosiddetto *"problema del bestiame"* di Archimede, pubblicato da Lessing nel 1773 da un manoscritto nella biblioteca di Wolfenbüttel [3]. Si afferma che il problema fu proposto da Archimede ad Eratostene, e la maggior parte degli esperti che si sono occupati della questione hanno raggiunto la conclusione che il problema fosse stato in effetti inventato da Archimede. Esso contiene otto incognite (numeri di capi di bestiame di vario tipo) che soddisfano sette equazioni lineari, assieme a due condizioni che asseriscono che certi numeri sono quadrati perfetti. Dopo alcuni passaggi di algebra elementare, il problema si riduce a risolvere l' equazione

$$t^2 - 472949u^2 = 1,$$

[1] L' equazione non presenta interesse quando d è un quadrato perfetto, dal momento che la differenza di due quadrati perfetti non può mai essere 1, tranne nel caso $1^2 - 0^2$.

la cui soluzione minima (esibita da Amthor nel 1880) è un numero naturale di quarantuno cifre, mentre la più piccola soluzione del problema originario consiste di numeri con centinaia di migliaia di cifre. Non vi è evidenza che gli antichi sapessero risolvere il problema, ma il solo fatto che l'avessero proposto suggerisce che essi potessero ben avere qualche nozione sull'equazione di Pell che sia poi andata smarrita.
Nei tempi moderni, il primo metodo sistematico per risolvere l'equazione fu esibito da Lord Brouncker nel 1657. Si tratta essenzialmente di sviluppare $\sqrt{d}$ in frazione continua, come verrà spiegato in seguito. Quasi contemporaneamente, Frènicle de Bessy (in un lavoro che non è pervenuto ai giorni nostri) tabulò soluzioni della

$$x^2 - dy^2 = 1 \qquad (\text{ o, equivalentemente } \quad x^2 = dy^2 + 1)$$

per tutti i valori di d fino a 150 e sfidò Brounecker a risolvere l'equazione $x^2 - 313y^2 = 1$. Brounecker, in risposta, produsse una soluzione (in cui x ha sessanta cifre), che disse di aver trovato col proprio metodo in una o due ore. Sia Wallis [6], durante l' esposizione del metodo di Brounecker, che Fermat, nel commentare il lavoro di Wallis, affermarono di aver dimostrato che l'equazione (0.0.1) è sempre risolubile. Fermat sembra essere stato il primo ad enunciare categoricamente che le soluzioni sono infinite. La prima dimostrazione di questo fatto, però, si deve a Lagrange, ed apparì intorno al 1766. Il nome di Pell fu associato all'equazione da Eulero per errore, egli pensò che il metodo di soluzione esibito da Wallis fosse dovuto a John Pell, un matematico inglese dello stesso periodo.
Le principali applicazioni dell'equazione di Pell le ritroviamo nella teoria delle forme quadratiche e nei problemi riguardanti il campo degli irrazionali quadratici; se $d \not\equiv 1 \ (mod 4)$ e $\alpha = x + y\sqrt{d}$, con $x, y \in \mathbb{N}$, allora α è un'unità in $\mathbb{Z}[\sqrt{d}]$ se e solo se $x^2 - dy^2 = \pm 1$; un risultato simile si ottiene se $d \equiv 1 \ (mod 4)$.
Questa tesi è essenzialmente un lavoro di approfondimento di alcuni argomenti inerenti alle frazioni continue e collegati allo studio delle soluzioni

dell'equazione di Pell. In particolare ci occuperemo del calcolo dei valori numerici interi x e y soddisfacenti l'equazione $x^2 - dy^2 = \pm 1$. Queste equazioni rientrano nel vasto ed interessante campo delle equazioni diofantee che sono equazioni da risolversi in interi (o in numeri razionali).

Il lavoro si articola essenzialmente in due sezioni principali.

Nella prima parte, introduttiva, daremo preliminarmente delle informazioni sulle frazioni continue; continueremo quindi con alcune osservazioni sullo sviluppo di un irrazionale quadratico, illustrando il metodo per il calcolo dei quozienti parziali a_i e per quello dei numeratori p_i e dei denominatori q_i delle ridotte delle frazioni continue. Svilupperemo inoltre $\sqrt{d}$, essendo d un intero non quadrato perfetto.
Nella seconda parte viene discussa l'equazione di Pell ed i suoi legami con i campi quadratici.

Capitolo 1

Richiami sulle frazioni continue

1.1 Frazioni continue generali e loro convergenti

Prima di cominciare a parlare dell' *equazione di Pell* è bene fare alcuni richiami sul concetto di frazione continua che più avanti potranno esserci utili.

Definizione 1.1.1. Una *frazione continua* [5] è un'espressione della forma

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots}}}$$

con un numero finito o infinito di termini $a_i \in \mathbb{Z}$. Un altro tipo di notazione per indicare la frazione continua è la seguente:

$$[a_0, a_1, a_2, \ldots].$$

Sono dette frazioni continue *semplici* quelle in cui gli a_i sono numeri interi positivi $\forall i > 0$. Mentre una frazione continua infinita $[a_0, a_1, a_2, \ldots]$ è detta *convergente* quando la sequenza di frazioni continue finite $[a_0]$, $[a_0, a_1]$, $[a_0, a_1, a_2]$, ... converge, (in generale, una frazione continua $[a_0, a_1, a_2, \ldots]$ converge se e solo se $\sum_{i=0}^{\infty} a_i$ diverge).

Definizione 1.1.2. Data una frazione continua $[a_0, a_1, \ldots, a_n]$ le quantità

$$\frac{p_0}{q_0} = a_0, \quad \frac{p_1}{q_1} = a_0 + \frac{1}{a_1}, \quad \frac{p_2}{q_2} = a_0 + \frac{1}{a_1 + \frac{1}{a_2}}, \quad \ldots$$

sono detti *convergenti* [8]. I numeri a_i sono detti *quozienti parziali*. L'ultimo convergente $\frac{p_n}{q_n}$ è uguale alla frazione stessa.

In generale, abbiamo il seguente teorema:

Teorema 1.1.1. *Per i convergenti* $\frac{p_n}{q_n}$ *della frazione continua* $[a_0, a_1, \ldots, a_n]$ *risulta:*

$$(i)\ p_n = [a_0, a_1, \ldots\ a_{n-1}, a_n] = a_n\,[a_0, \ldots, a_{n-1}] + [a_0, \ldots, a_{n-2}] = a_n p_{n-1} + p_{n-2}, \tag{1.1.1}$$

per $(n \geq 2)$

$$(ii)\ q_n = [a_1, a_2, \ldots\ a_{n-1}, a_n] = a_n\,[a_1, \ldots, a_{n-1}] + [a_1, \ldots, a_{n-2}] = a_n q_{n-1} + q_{n-2}, \tag{1.1.2}$$

per $(n \geq 2)$

Dimostrazione. (i) e (ii) Per i valori iniziali si ha che:

$$p_0 := [a_0] = a_0; \qquad p_1 := [a_0, a_1] = a_0 a_1 + 1;$$

$$q_0 := 1; \qquad q_1 := [a_1] = a_1;$$

Per induzione sia $p_i = [a_0, a_1, \ldots, a_{i-1}, a_i]$ e $q_i = [a_1, a_2, \ldots, a_{i-1}, a_i]$ per ogni $i < n$.
Dalla definizione di frazione continua si ha che:

$$\frac{p_n}{q_n} = [a_0, a_1, \ldots, a_n] = a_0 + \frac{1}{[a_1, a_2, \ldots, a_n]} = a_0 + \frac{1}{\frac{[a_1, a_2, \ldots, a_n]}{[a_2, a_3, \ldots, a_n]}} =$$

$$a_0 + \frac{[a_2, a_3, \ldots, a_n]}{[a_1, a_2, \ldots, a_n]} = \frac{a_0\,[a_1, a_2, \ldots, a_n] + [a_2, a_3, \ldots, a_n]}{[a_1, a_2, \ldots, a_n]} = \frac{[a_0, a_1, \ldots, a_{n-1}, a_n]}{[a_1, a_2, \ldots, a_{n-1}, a_n]} \tag{1.1.3}$$

La frazione continua generale ha dunque un valore dato da:

$$\frac{p_n}{q_n} = [a_0, a_1, \ldots, a_n] = a_0 + \frac{1}{a_1+}\frac{1}{a_2+}\cdots\frac{1}{a_n} = \frac{[a_0,a_1,\ldots,a_{n-1},a_n]}{[a_1,a_2,\ldots,a_{n-1},a_n]}$$

Tornando alla penultima uguaglianza della (1.1.3) si ha che:

$$\frac{p_n}{q_n} = \frac{a_0[a_1,a_2,\ldots,a_n]+[a_2,a_3,\ldots,a_n]}{[a_1,a_2,\ldots,a_n]} = \frac{a_n[a_{n-1},a_{n-2},\ldots,a_0]+[a_{n-2},a_{n-3},\ldots,a_0]}{[a_n,a_{n-2},\ldots,a_1]} = \frac{a_n p_{n-1}+p_{n-2}}{a_n q_{n-1}+q_{n-2}}.$$

□

Lemma 1.1.1. ([3] , p. 76) *Qualsiasi coppia di convergenti consecutivi, nello sviluppo in frazioni continue di un numero razionale, soddisfa la relazione:*

$$p_m q_{m-1} - p_{m-1} q_m = (-1)^{m-1} \tag{1.1.4}$$

Dimostrazione. Sia m=1 allora:

$$p_0 := [a_0] = a_0; \qquad p_1 := [a_0, a_1] = a_0 a_1 + 1;$$

$$q_0 := 1; \qquad q_1 := [a_1] = a_1;$$

Pertanto

$$p_1 q_0 - p_0 q_1 = (a_0 a_1 + 1) \cdot 1 - a_0 a_1 = 1 \tag{1.1.5}$$

Usando le relazioni (1.1.1) e (1.1.2) si ottiene che:

$$p_m q_{m-1} - p_{m-1} q_m = (a_m p_{m-1} + p_{m-2}) q_{m-1} - p_{m-1}(a_m q_{m-1} + q_{m-2}) = -(p_{m-1} q_{m-2} - p_{m-2} q_{m-1}).$$

Pertanto se chiamo il primo membro della relazione (1.1.4) Δ_m si ha che:

$$\Delta_m = -\Delta_{m-1}$$

e continuando

$$\Delta_m = -\Delta_{m-1} = +\Delta_{m-2} = \ldots = \pm\Delta_1$$

dove il segno alla fine è $+1$ se m è dispari e -1 se è pari, cosicché lo si può rappresentare con

$$\Delta_m = (-1)^{m-1}$$

Dal momento che $\Delta_1 = 1$, in virtù della (1.1.5), il risultato generale (1.1.4) resta dimostrato.

□

1.2 Frazioni continue puramente periodiche

Definizione 1.2.1. Un numero irrazionale si dice *quadratico* ([3], p. 84), se è soluzione di una certa equazione quadratica con coefficienti interi. In particolare, la radice quadrata di qualsiasi numero naturale N che non sia un quadrato perfetto, è un irrazionale quadratico, in quanto soluzione dell'equazione

$$x^2 - N = 0.$$

Definizione 1.2.2. Si definisce *quoziente completo* ([3], p. 90) una relazione della forma

$$\alpha_n = \frac{p_n+\sqrt{d}}{q_n},$$

dove p_n e q_n sono numeri naturali che soddisfano

$$p_n < \sqrt{d}, \qquad q_n < 2\sqrt{d},$$

e hanno la proprietà che $p_n^2 - d$ sia multiplo di q_n.

Definizione 1.2.3. Una *frazione continua periodica* è una frazione continua i cui termini si ripetono da un certo punto in poi. Il numero più piccolo di termini ripetuti è chiamato periodo della frazione continua. Ogni frazione continua periodica rappresenta un numero irrazionale.
La radice quadrata di un intero ha una frazione continua periodica della forma:

$$\sqrt{d} = \left[a_0, \overline{a_1, \ldots, 2a_0}\right]$$

([5], p. 130), dove la parte che si ripete (escluso l'ultimo termine) è simmetrica e il termine centrale appare o una volta o due.

Definizione 1.2.4. Data una frazione continua periodica $\sqrt{d} = \left[a_0, \overline{a_1, \ldots, 2a_0}\right]$ si dirà che $\{a_0\}$ è l'*antiperiodo* e $\{a_1, \ldots, 2a_0\}$ il *periodo*.
Una frazione continua periodica si dice *periodica pura* se il suo antiperiodo è vuoto.

Iniziamo ora a trattare frazioni continue periodiche enunciando il seguente teorema attribuito a Galois (1828):

Teorema 1.2.1. *Una frazione continua semplice è puramente periodica se e solo se:*

$$(i) \quad \alpha \text{ un numero reale algebrico di grado } 2 \text{ su } \mathbb{Q};$$

$$(ii) \quad \alpha > 1;$$

$$(iii) \quad \text{il coniugato di } \alpha \text{ soddisfa la disuguaglianza } -1 < \alpha' < 0.$$

In particolare se α soddisfa le (i), (ii), (iii) si dice ridotto.

Dimostrazione. $''\Rightarrow''$Sia $\alpha = [\overline{a_0, a_1, a_2, \ldots, a_{n-1}, a_n}]$. Se il periodo inizia con a_0 allora $a_0 = a_{n+1} \geq 1$, e cioè $\alpha > 1$ (ii).
Inoltre dall'equazione generale

$$\alpha = \frac{p}{q} = \frac{\alpha_{n+1} p_n + p_{n-1}}{\alpha_{n+1} q_n + q_{n-1}} \tag{1.2.1}$$

segue che

$$q_n \alpha^2 - (p_n - q_{n-1})\alpha - p_{n-1} = 0$$

essendo $\alpha_{n+1} = \alpha$. Il polinomio di secondo grado

$$f(x) = q_n x^2 - (p_n - q_{n-1})x - p_{n-1} = 0$$

è irriducibile in $\mathbb{Q}[x]$ (i).
Infine per la regola di Eulero si ha:

$$p_n = [a_0, a_1, \ldots, a_{n-1}, a_n] \qquad q_n = [a_1, a_2, \ldots, a_{n-1}, a_n]$$

Consideriamo ora la frazione continua che si ottiene da α rovesciando il periodo:

$$\beta = [\overline{a_n, a_{n-1}, \ldots, a_0}]$$

Notiamo che β è maggiore di 1 essendo $a_n \geq 1$ e dall'equazione generale si ha:

$$\beta = \frac{\beta_{n+1}p_n+q_n}{\beta_{n+1}p_{n-1}+q_{n-1}} = \frac{\beta p_n+q_n}{\beta p_{n-1}+q_{n-1}}$$

da cui

$$p_{n-1}\beta^2 - (p_n - q_{n-1})\beta - q_n = 0$$

che è equivalente all'equazione:

$$q_n(-\tfrac{1}{\beta})^2 - (p_n - q_{n-1})(-\tfrac{1}{\beta}) - p_{n-1} = 0$$

Allora $-\frac{1}{\beta}$ è zero di $f(x)$ diverso da α e poiché $\beta > 1$ si ha che $-1 < -\frac{1}{\beta} < 0$, cioè $-\frac{1}{\beta}(= \alpha')$ soddisfa la (iii).

$''\Leftarrow''$Sia α zero reale positivo del polinomio $f(x) = ax^2+bx+c$ irriducibile in $\mathbb{Z}[x]$; sia poi α' l'altro zero di f(x) e si supponga che $-1 < \alpha' < 0$. Quindi

$$\alpha = \frac{-b+\sqrt{b^2-4ac}}{2a} = \frac{P+\sqrt{D}}{Q}$$
$$\alpha' = \frac{-b-\sqrt{b^2-4ac}}{2a} = \frac{P-\sqrt{D}}{Q}$$

dove $P, Q \in \mathbb{Z}$, e D intero positivo (non-quadrato). Per ipotesi α è ridotto, allora $\alpha > 1$ e $-1 < \alpha' < 0$, pertanto:

1. $\alpha - \alpha' > 0 \;\Rightarrow\; \frac{2\sqrt{D}}{Q} \;\Rightarrow\; Q > 0$;
2. $\alpha + \alpha' > 0 \;\Rightarrow\; \frac{2P}{Q} > 0 \;\Rightarrow\; P > 0$;
3. $\alpha' < 0 \;\Rightarrow\; \frac{P-\sqrt{D}}{Q} < 0 \;\Rightarrow\; P < \sqrt{D}$;

4. $\alpha > 1 \;\Rightarrow\; \frac{P+\sqrt{D}}{Q} > 1 \;\Rightarrow\; P + \sqrt{D} > Q$;

Riassumendo:

$$0 < P < \sqrt{D} \quad e \quad 0 < Q < 2\sqrt{D} \tag{1.2.2}$$

Inoltre poiché $Q = \pm 2a$ e $P^2 - D = (-b)^2 - (b^2 - 4ac)$ si ha che:

$$Q | P^2 - D \tag{1.2.3}$$

Ricordiamo le proprietà dei coniugati:

$$(a_1 + a_2)' = a_1' + a_2'; \quad (a_1 - a_2)' = a_1' - a_2'; \quad (a_1 a_2)' = a_1' a_2'; \quad (\frac{a_1}{a_2})' = \frac{a_1'}{a_2'}. \tag{1.2.4}$$

Possiamo ora procedere con la dimostrazione.
Sia $\alpha = a_0 + \frac{1}{\alpha_1}$ dove $a_0 (\geq 1)$ è la parte intera di α e $\frac{1}{\alpha_1}$ è la parte frazionaria di α (con $\alpha_1 > 1$). Vediamo che anche α_1 è un irrazionale quadratico ridotto e applicando ad α le proprietà dei coniugati (1.2.4) si ottiene:

$$(\alpha = a_0 + \tfrac{1}{\alpha_1})'$$
$$\alpha' = a_0 + \tfrac{1}{\alpha_1'}$$

da cui

$$\alpha_1' = -\tfrac{1}{a_0 - \alpha'} \text{ dove } -1 < \alpha' < 0$$

dunque $-1 < \alpha_1' < 0$, cioé α_1 è ridotto. Analogamente anche $\alpha_2, \alpha_3, \ldots, \alpha_n, \ldots$ sono irrazionali quadratici ridotti. Per quanto riguarda la forma di α_1 si ha

$$\tfrac{1}{\alpha_1} = \alpha - a_0 = \tfrac{P+\sqrt{D}}{Q} - a_0 = \tfrac{P - Qa_0 + \sqrt{D}}{Q}$$

da cui

$$\alpha_1 = \frac{Q}{P - Qa_0 + \sqrt{D}}$$

Sia

$$P_1 = -P + Qa_0. \tag{1.2.5}$$

Allora

$$\alpha_1 = \frac{Q}{-P_1 + \sqrt{D}} = \frac{Q(\sqrt{D} + P_1)}{D - P_1^2}.$$

Ora poniamo

$$Q_1 = \frac{D - P_1^2}{Q} \tag{1.2.6}$$

da cui si ottiene

$$\alpha_1 = \frac{P_1 + \sqrt{D}}{Q_1} \tag{1.2.7}$$

Osserviamo che Q_1 è intero, in quanto $P^2 - D$ è multiplo di Q e $P_1 \equiv -P \ (mod Q)$.
Inoltre essendo α_1 ridotto, gli interi P_1 e Q_1 sono positivi e soddisfano le condizioni (1.2.2) e $P_1^2 - D$ è multiplo di Q_1, in virtù della (1.2.6).
Quindi possiamo ripetere il ragionamento fatto finora partendo con α_1 al posto di α e tutto funziona. In generale ogni quoziente completo ha la forma:

$$\alpha_n = \frac{P_n + \sqrt{D}}{Q_n}$$

dove P_n e Q_n sono interi positivi che soddisfano la (1.2.2).
Vi sono, per la (1.2.2), solo un numero finito di possibilità per P_n e Q_n, e prima o poi ci imbatteremo in un paio di valori già incontrato in precedenza. Ovvero troveremo un quoziente completo uguale a uno precedente, e da quel punto la frazione continua sarà *periodica*.
Dopo un certo numero di passi si ottiene una ripetizione della coppia $(P_r, Q_r) =$

(P_{r+n+1}, Q_{r+n+1}) e quindi la periodicità $\alpha_r = \alpha_{r+n+1}$.
Dobbiamo ancora dimostrare che la frazione continua è *puramente periodica*, ossia periodica fin dall'inizio.
Introduciamo, per ogni i, il numero:

$$\beta_i = -\tfrac{1}{\alpha_i'} \ , \ (-1 < \alpha_i' < 0 \ \Rightarrow \ -\tfrac{1}{\alpha_i'} = \beta_i > 1)$$

e coniugando α_i si ottiene

$$(\alpha_i = a_i + \tfrac{1}{\alpha_{i+1}})'$$

$$\alpha_i' = a_i + \frac{1}{\alpha_{i+1}'}. \tag{1.2.8}$$

L'ultima relazione prende la forma

$$-\frac{1}{\beta_n} = a_n - \beta_{n+1} \ \ o \ \ \beta_{n+1} = a_n + \frac{1}{\beta_n}. \tag{1.2.9}$$

Osservando le relazioni (1.2.8) e (1.2.9) notiamo che: la parte intera di α_i' cioè

$[\alpha_i'] = a_i = [\beta_{i+1}]$ essendo $\alpha_r = \alpha_{r+n+1} \ \Rightarrow \ \alpha_r' = \alpha_{r+n+1}' \ \Rightarrow \ \beta_r = \beta_{r+n+1}$

ma $([\beta_r] = a_{r-1}$ e $[\beta_{r+n+1}] = a_{r-1+n+1}) \ \Rightarrow \ a_{r-1} = a_{r+n}$.
Ora $(a_{r-1} = a_{r+n}$ e $\alpha_r = \alpha_{r+n+1}) \ \Rightarrow \ \alpha_{r-1} = \alpha_{r-1+n+1}$ (poiché $\alpha_{r-1} = a_{r-1} + \frac{1}{\alpha_{r-1}}$ e $\alpha_{r-1+n+1} = a_{r-1+n+1} + \frac{1}{\alpha_{r-1+n+1}}$).
Iterando il procedimento si trova che: $\alpha_0 = \alpha_{n+1}$ come si voleva. Quindi α è puramente periodica.

□

Capitolo 2

Equazione di Pell

Sia d un numero naturale non quadrato perfetto. L'equazione diofantea che consideriamo è:

$$x^2 - dy^2 = 1 \qquad o \qquad x^2 = dy^2 + 1 \tag{2.0.1}$$

2.1 Soluzioni dell'equazione di Pell

Vediamo ora come trovare le soluzioni dell'equazione (2.0.1).

Proposizione 2.1.1. *Se* (x_1, y_1) *è una soluzione di*

$$x^2 - dy^2 = \pm 1 \tag{2.1.1}$$

allora $\frac{x_1}{y_1}$ *è un convergente dello sviluppo in frazioni continue di* $\sqrt{d}$*, con* $\sqrt{d} = \left[a_0, \overline{a_1, \ldots, 2a_0}\right]$*, cioè* $\frac{x_1}{y_1} = \frac{p_i}{q_i}$ *per qualche* i*.*

Dimostrazione. ([9], th. 184) Per ipotesi si ha che:

$$x_1^2 - dy_1^2 = \pm 1$$

cioè

$$(x_1 + y_1\sqrt{d})(x_1 - y_1\sqrt{d}) = \pm 1$$

moltiplico entrambi i membri per $\frac{1}{x_1+y_1\sqrt{d}}$ ed ottengo:

$$(x_1 - y_1\sqrt{d}) = \pm\frac{1}{x_1+y_1\sqrt{d}}$$

$$y_1(\frac{x_1}{y_1} - \sqrt{d}) = \pm\frac{1}{y_1(\frac{x_1}{y_1}+\sqrt{d})}$$

divido ancora entrambi i membri per y_1 ed ottengo:

$$|\frac{x_1}{y_1} - \sqrt{d}| = \frac{1}{y_1^2(\frac{x_1}{y_1}+\sqrt{d})} < \frac{1}{2\sqrt{d}y_1^2} < \frac{1}{2y_1^2};$$

perciò

$$|\frac{x_1}{y_1} - \sqrt{d}| < \frac{1}{2y_1^2}. \tag{2.1.2}$$

□

Consideriamo il problema inverso: per quali i, $\frac{p_i}{q_i}$ è una soluzione dell'equazione (2.0.1)?

Consideriamo la frazione continua di $\sqrt{d}$

$$\sqrt{d} = \left[a_0, \overline{a_1, \ldots, 2a_0}\right] = a_0 + \frac{1}{a_1+}\frac{1}{a_2+}\cdots\frac{1}{a_n+}\frac{1}{2a_0+}\frac{1}{a_1+}\cdots$$

Sia ora

$$\frac{p_{n-1}}{q_{n-1}} = a_0 + \frac{1}{a_1+}\frac{1}{a_2+}\cdots\frac{1}{a_{n-1}}$$

$$\frac{p_n}{q_n} = a_0 + \frac{1}{a_1+}\frac{1}{a_2+}\cdots\frac{1}{a_{n-1}+}\frac{1}{a_n};$$

il quoziente completo

$$\alpha_{n+1} = 2a_0 + \frac{1}{a_1+}\frac{1}{a_2+}\cdots\frac{1}{a_{n-1}+}\frac{1}{a_n+}\frac{1}{2a_0+}\frac{1}{a_1+}\cdots = \sqrt{d} + a_0 \tag{2.1.3}$$

questo perché la frazione continua per $\sqrt{d}$ non può essere puramente periodica, poiché il coniugato di $\sqrt{d}$ è $-\sqrt{d}$, e questo non è compreso tra -1 e 0. Consideriamo invece il numero $\sqrt{d}+a_0$, dove a_0 è la parte intera di $\sqrt{d}$. Il coniugato di questo numero è $-\sqrt{d}+a_0$, che giace tra -1 e 0. Dunque la frazione continua per $\sqrt{d}+a_0$ è puramente periodica, e poiché ha inizio con $2a_0$, essa ha la forma (2.1.3).
Dall'equazione (1.2.1) si ha:

$$\sqrt{d} = \frac{\alpha_{n+1}p_n + p_{n-1}}{\alpha_{n+1}q_n + q_{n-1}} = \frac{(\sqrt{d}+a_0)p_n + p_{n-1}}{(\sqrt{d}+a_0)q_n + q_{n-1}}$$

da cui:

$$\sqrt{d}((\sqrt{d}+a_0)q_n + q_{n-1}) = (\sqrt{d}+a_0)p_n + p_{n-1}$$

$$dq_n + (a_0q_n + q_{n-1})\sqrt{d} = (a_0p_n + p_{n-1}) + p_n\sqrt{d}$$

Quest'ultima è un'equazione del tipo $a+b\sqrt{d} = c+f\sqrt{d}$ dove $a, b, c, f \in \mathbb{Z}$ e $\sqrt{d}$ è irrazionale. Perciò si avrà che a=c e b=f, cioè:

$$\begin{cases} dq_n = a_0p_n + p_{n-1} \\ a_0q_n + q_{n-1} = p_n \end{cases}$$

Risolvendo per p_{n-1} e q_{n-1} abbiamo che:

$$\begin{cases} p_{n-1} = dq_n - a_0p_n \\ q_{n-1} = p_n - a_0q_n \end{cases}$$

Noi sappiamo dalla (1.1.4) che

$$p_mq_{m-1} - p_{m-1}q_m = (-1)^{m-1} \tag{2.1.4}$$

quindi sostituendo:

$$p_n(p_n - a_0q_n) - (dq_n - a_0p_n)q_n = (-1)^{n-1}$$

da cui si ottiene l'equazione:

$$p_n^2 - dq_n^2 = (-1)^{n-1} \tag{2.1.5}$$

A questo punto si verificano 2 casi:

1. Se $n-1$ è pari, cioè n è dispari $\Rightarrow p_n^2 - dq_n^2 = 1$
2. Se $n-1$ è dispari, cioè n è pari $\Rightarrow p_n^2 - dq_n^2 = -1$

Nel primo caso (n dispari) abbiamo una soluzione dell'equazione di Pell (2.0.1) e cioè:

$$x = p_n \qquad y = q_n$$

dove $\frac{p_n}{q_n}$ sono i convergenti nello sviluppo in frazioni continue di $\sqrt{d}$ troncati prima del termine $2a_0$.

Nel secondo caso (n pari) abbiamo una soluzione dell'equazione

$$x^2 - dy^2 = -1 \tag{2.1.6}$$

dal momento che $(-1)^{n-1}$ è negativo, ma

$$p_{2n+1}^2 - dq_{2n+1}^2 = 1$$

perciò sostituiamo n con $2n+1$ nell'equazione (2.1.5):

$$p_n^2 - dq_n^2 = (-1)^{2n} = 1$$

ottenendo così una soluzione dell'equazione di Pell (2.0.1) e cioè:

$$x = p_{2n+1} \qquad y = q_{2n+1}$$

dove p_{2n+1} e q_{2n+1} sono i quozienti completi nello sviluppo in frazioni continue di $\sqrt{d}$ troncati prima del termine $2a_0$ che incontro per la seconda volta.

Riassumendo:

$$(x, y) = \begin{cases} (p_n, q_n) & \text{se } n \text{ è dispari} \\ (p_{2n+1}, q_{2n+1}) & \text{se } n \text{ è pari} \end{cases}$$

Analogamente è possibile trovare soluzioni dell'equazione (2.1.6).
Infatti può essere risolta a partire dall'equazione (2.0.1) se n è pari, ma non ha soluzioni se n è dispari, ossia:

$$(x, y) = \begin{cases} (p_n, q_n) & \text{se } n \text{ è dispari} \\ \textit{nessuna soluzione} & \text{se } n \text{ è pari} \end{cases}$$

In definitiva per mostrare che l'equazione di Pell ha infinite soluzioni e che queste sono ottenute a partire dai convergenti di $\sqrt{d}$ che corrispondono ai termini a_n alla fine di ogni periodo, si procede nel seguente modo:

1. se n è dispari (cioè la frazione continua ha un termine centrale) tutti questi sono soluzioni dell'equazione di Pell (2.0.1);

2. se n è pari (cioè la frazione continua non ha un termine centrale) i convergenti forniscono alternativamente soluzioni dell'equazione con -1 (2.1.6) e dell'equazione con 1 (2.0.1).

   Pertanto le soluzioni sono infinite.

Nel caso in cui il periodo di $\left[a_0, \overline{a_1, \ldots, a_1, 2a_0}\right]$ ha lunghezza $n+1$ pari, n è dispari (cioè palindromo $a_1, a_2, \ldots, a_2, a_1$ ha un elemento centrale) allora la (2.1.6) non ha soluzioni, mentre la (2.0.1) è risolta dalle coppie

$$(p_n, q_n); \qquad (p_{2n}, q_{2n}); \qquad (p_{3n}, q_{3n}); \qquad \ldots$$

Nel caso in cui il periodo di $\left[a_0, \overline{a_1, \ldots, a_1, 2a_0}\right]$ ha lunghezza $n+1$ dispari, n è pari (cioè palindromo $a_1, a_2, \ldots, a_2, a_1$ non ha un elemento centrale) allora la (2.1.6) è risolta dalle coppie

$$(p_n, q_n); \qquad (p_{2n}, q_{2n}); \qquad (p_{3n}, q_{3n}); \qquad \ldots$$

mentre la (2.0.1) è risolta dalle coppie

$$(p_{2n+1}, q_{2n+1}); \qquad (p_{4n+3}, q_{4n+3}); \qquad (p_{6n+5}, q_{6n+5}); \qquad \ldots$$

La distinzione nei casi in cui n è dispari o pari solleva problemi ai quali non si è finora data una risposta completa. Infatti non è ancora noto il modo di caratterizzare i numeri d per cui n è pari [1].

Data una soluzione (x,y)=(p,q), che può essere trovata come sopra, tutto l'insieme delle soluzioni può essere dedotto elevando tutto all'r-esima potenza, ossia:

$$x_r^2 - dy_r^2 = (p^2 - dq^2)^r = 1.$$

Fattorizzando abbiamo

$$(x_r + \sqrt{d}y_r)(x_r - \sqrt{d}y_r) = (p + \sqrt{d}q)^r(p - \sqrt{d}q)^r$$

cioè

$$(x_r + \sqrt{d}y_r) = (p + \sqrt{d}q)^r$$

$$(x_r - \sqrt{d}y_r) = (p - \sqrt{d}q)^r$$

che restituisce l'insieme di soluzioni:

$$x_r = \frac{(p+\sqrt{d}q)^r + (p-\sqrt{d}q)^r}{2}$$

$$y_r = \frac{(p+\sqrt{d}q)^r - (p-\sqrt{d}q)^r}{2\sqrt{d}}$$

(soluzioni dell'equazione (2.0.1) sia nel caso in cui r sia pari sia nel caso in cui r sia dispari).
Queste soluzioni possono essere utizzate anche per la risoluzione dell'equazione (2.1.6) ma solo nel caso in cui r sia dispari.

Concludendo l'equazione di Pell (2.0.1) può sempre essere risolta, mentre l'equazione (2.1.6) non è sempre risolubile.

Seguiamo il seguente ragionamento. Se l'equazione (2.1.6)

$$x^2 - dy^2 = -1$$

è risolubile, allora è risolubile la congruenza

$$x^2 + 1 \equiv 0 \ (mod \ d) \qquad \Leftrightarrow \qquad x^2 \equiv -1 \ (mod \ d)$$

Questo significa che -1 è un quadrato in $\mathbb{Z}/d\mathbb{Z}$.

Se p è un numero primo diverso da 2, andiamo allora a considerare il simbolo di Legendre:

$$\left(\frac{-1}{p}\right) = \begin{cases} 1 & \text{se -1 è un quadrato in } \mathbb{Z}/p\mathbb{Z} \text{ (cioè } p \equiv 1 \pmod 4) \\ -1 & \text{se -1 è un non quadrato in } \mathbb{Z}/p\mathbb{Z} \text{ (cioè } p \equiv 3 \pmod 4) \end{cases}$$

Nel nostro caso, se $p \mid d$ da $x^2 \equiv -1 \pmod d$ segue che

$$x^2 \equiv -1 \pmod p \qquad \Leftrightarrow \qquad \left(\tfrac{-1}{p}\right) = 1 \qquad \Leftrightarrow \qquad p \equiv 1 \pmod 4$$

Quindi una condizione necessaria affinché la (2.1.6) sia risolubile è che ogni fattore primo di d sia della forma 4k+1 e che d non sia divisibile per 4. Però tale condizione non è sufficiente per l'esistenza di una soluzione, come dimostrato dall'equazione $x^2 - 34y^2 = -1$, che non ha soluzioni in interi ([10], p. 201 e 204).

2.2 Genesi moltiplicativa delle soluzioni dell'equazione di Pell

Continuiamo a vedere come ottenere soluzioni dell'equazione

$$x_1^2 - dy_1^2 = 1$$

Proposizione 2.2.1. *([4], p. 277) Se α è un numero irrazionale allora esistono infiniti numeri razionali x/y, $(x, y) = 1$ tali che $|x/y - \alpha| < 1/y^2$.*

Dimostrazione. Una partizione dell'intervallo semiaperto $[0, 1)$ è data da

$$[0,1) = \left[0, \frac{1}{n}\right) \cup \left[\frac{1}{n}, \frac{2}{n}\right) \cup \ldots \cup \left[\frac{n-1}{n}, 1\right).$$

Se indichiamo con $[\eta]$ la parte intera superiore di η allora la sua parte frazionaria è data da $\eta - [\eta]$. Essa si trova in un unico termine della partizione. Consideriamo le parti frazionarie di $0, \alpha, 2\alpha, \ldots, n\alpha$. Almeno due di queste devono trovarsi nello stesso sottointervallo. Devono esistere, cioè, j, k con $j > k$, $0 \leq j, k \leq n$ tali che

$$|j\alpha - [j\alpha] - (k\alpha - [k\alpha])| < \frac{1}{n}. \tag{2.2.1}$$

Sia $y = j-k$, $x = [k\alpha]-[j\alpha]$ tali che l'espessione (2.2.1) diventa $|x-y\alpha| < 1/n$. Assumiamo che x e y siano coprimi. Da $0 < y < n$ risulta $|x/y - \alpha| < 1/ny < 1/y^2$. Per ottenere infinite soluzioni notiamo che $|x/y - \alpha| \neq 0$ e scegliamo un intero $m > 1/|x/y - \alpha|$. Il procedimento descritto garantisce l'esistenza di interi x_1, y_1 tali che $|x_1/y_1 - \alpha| < 1/my_1 < |x/y - \alpha|$ e $0 < y_1 < m$. Questo procedimento fornisce un numero infinito di soluzioni.

□

Questa proposizione sarà applicata per dimostrare che $|x^2 - dy^2|$ assume lo stesso valore infinite volte.

Lemma 2.2.1. *([4], p. 277) Se d è un numero intero positivo (non-quadrato) allora esiste una costante M tale che $|x_n^2 - dy_n^2| < M$ ha infinite soluzioni intere.*

Dimostrazione. Scriviamo $x_n^2 - dy_n^2 = (x_n + \sqrt{d}y_n)(x_n - \sqrt{d}y_n)$. Abbiamo visto in precedenza che esistono infinite coppie di numeri coprimi (x_n, y_n), $y_n > 0$ che soddisfano $|x_n - \sqrt{d}y_n| < \frac{1}{y_n}$. Da questo segue che $|x_n + \sqrt{d}y_n| < |x_n - \sqrt{d}y_n| + 2\sqrt{d}|y_n| < \frac{1}{y_n} + 2\sqrt{d}y_n$. Da ciò $|x_n^2 - dy_n^2| < \frac{1}{y_n}|\frac{1}{y_n} + 2\sqrt{d}y_n| \leq 2\sqrt{d} + 1$ □

Teorema 2.2.1. *([4], p. 277) Se d è un numero intero positivo (non-quadrato) allora $x_n^2 - dy_n^2 = 1$ ha infinite soluzioni intere. Inoltre se (x_1, y_1) è la più piccola soluzione positiva dell'equazione (2.0.1) $x^2 - dy^2 = 1$, cioè la soluzione per cui è minimo il numero $x_1 + y_1\sqrt{d}$, allora tutte le altre soluzioni $\pm(x_n, y_n)$ possono essere ottenute dall'equazione:*

$$x_n + y_n\sqrt{d} = (x_1 + y_1\sqrt{d})^n \qquad dove \quad n = 1, 2, 3, \ldots \tag{2.2.2}$$

Dimostrazione. Dal lemma (2.2.1) sappiamo che esiste $m \in \mathbb{Z}$ tale che $x_n^2 - dy_n^2 = m$ per infinite coppie (x_n, y_n), $x_n > 0$, $y_n > 0$. Assumiamo che le x_n componenti siano distinte e che vi sia un numero finito di classi resto modulo $|m|$. Per esempio siano (x_1, y_1), (x_2, y_2), $x_1 \neq x_2$ tali che $x_1 \equiv x_2 \quad (|m|)$, $y_1 \equiv y_2 \quad (|m|)$. Prendiamo $\alpha = x_1 - y_1\sqrt{d}$, $\beta = x_2 - y_2\sqrt{d}$. Se $\gamma = x_n - y_n\sqrt{d}$

denotiamo il coniugato di γ con $\gamma' = x_n + y_n\sqrt{d}$ e la sua norma con $N(\gamma) = x_n^2 - dy_n^2$. Ricordiamo che $N(\alpha\beta) = N(\alpha)N(\beta)$. Con un rapido calcolo si mostra che $\alpha\beta' = X + Y\sqrt{d}$ dove $m|X$, $m|Y$. Così $\alpha\beta' = m(u + v\sqrt{d})$ per qualche intero u e v. Prendendo la norma di entrambi i membri otteniamo $m^2 = m^2(u^2 - v^2 d)$, da qui

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

Resta da dimostrare che $v \neq 0$.
In qualsiasi caso se $v = 0$ allora $u = \pm 1$ e $\alpha\beta' = \pm m$. Moltiplicando entrambi i membri per β otteniamo $\alpha m = \pm m\beta$ ossia $\alpha = \pm\beta$; ma questo implica che $x_1 = x_2$.
Così l'equazione di Pell ha una soluzione con $x_n y_n \neq 0$.
Proviamo ora la seconda asserzione secondo la quale possiamo affermare che le (x_n, y_n) sono soluzioni. Per esempio se (x_1, y_1) è la più piccola soluzione positiva della (2.0.1) la successiva soluzione (x_2, y_2) è data da

$$x_2 + y_2\sqrt{d} = (x_1 + y_1\sqrt{d})^2 = x_1^2 + dy_1^2 + 2x_1y_1\sqrt{d}$$

cioè

$$x_2 = x_1^2 + dy_1^2 \qquad y_2 = 2x_1y_1.$$

Mostriamo che effettivamente (x_2, y_2) è soluzione:

$$\begin{aligned} x_2^2 - dy_2^2 = \quad & (x_1^2 + y_1^2 d)^2 - d(2x_1y_1)^2 \\ & = x_1^4 - 2dx_1^2y_1^2 + d^2y_1^4 \\ & = (x_1^2 + y_1^2 d)^2 = (1)^2 = 1 \end{aligned}$$

È semplice mostrare che se (x_n, y_n) sono calcolate tramite l'equazione (2.2.2) allora $x_n^2 - dy_n^2 = 1$. Dalla (2.2.2) si ha che

$$x_n + y_n\sqrt{d} = \underbrace{(x_1 + y_1\sqrt{d})(x_1 + y_1\sqrt{d})\ldots(x_1 + y_1\sqrt{d})}_{n}$$

Coniugando si ottiene:

$$x_n - y_n\sqrt{d} = \underbrace{(x_1 - y_1\sqrt{d})(x_1 - y_1\sqrt{d})\dots(x_1 - y_1\sqrt{d})}_{n}$$

e cioè

$$x_n + y_n\sqrt{d} = (x_1 + y_1\sqrt{d})^n.$$

Concludendo

$$\begin{aligned} x_n^2 - dy_n^2 = & \quad (x_n + y_n\sqrt{d})(x_n - y_n\sqrt{d}) \\ & = (x_1 + y_1\sqrt{d})^n(x_1 - y_1\sqrt{d})^n \\ & = (x_1^2 - y_1^2 d)^n = (1)^n = 1 \end{aligned}$$

Pertanto x_n e y_n sono soluzioni dell'equazione (2.0.1).

Proviamo ora, che le soluzioni sono tutte di questo tipo.
Supponiamo per assurdo che (a, b) sia una soluzione positiva dell'equazione $x^2 - dy^2 = 1$ non ottenuta dalla (2.2.2).
Allora per certi interi positivi n si ha che:

$$(x_1 + y_1\sqrt{d})^n < a + b\sqrt{d} < (x_1 + y_1\sqrt{d})^{n+1}.$$

Allora dividendo per $(x_1 + y_1\sqrt{d})^n$:

$$1 < (a + b\sqrt{d})(x_n - y_n\sqrt{d}) < x_1 + y_1\sqrt{d}$$

Ma se $(a + b\sqrt{d})(x_n - y_n\sqrt{d}) = X + Y\sqrt{d}$, (X, Y) è una soluzione dell'equazione di Pell e $1 < X + Y\sqrt{d} < x_1 + y_1\sqrt{d}$.
Ora $X + Y\sqrt{d} > 0$ e $0 < X - Y\sqrt{d} = (X + Y\sqrt{d})^{-1} < 1$ si ottiene che $X > 0$ e $Y > 0$. Questo però va contro il fatto che (x_1, y_1) sia la più piccola soluzione positiva di $x^2 - dy^2 = 1$.

Pertanto tutte le soluzioni sono date dall'equazione (2.2.2). □

2.3 Unità dei campi quadratici ed equazione di Pell

Con riferimento all'equazione (2.1.1) ($x^2 - dy^2 = \pm 1$) nel campo reale ($\mathbb{R}$) consideriamo il sottoanello $\mathbb{Z}\left[\sqrt{d}\right] = \left\{a + b\sqrt{d} | a, b \in \mathbb{Z}\right\}$, (essendo d un $non - quadrato$ in $\mathbb{Z}$) e studiamo il gruppo U dei suoi elementi invertibili. Ad ogni elemento di $\mathbb{Z}\left[\sqrt{d}\right]$ si associa una *norma* definita come segue.

$$N(a + b\sqrt{d}) := a^2 - db^2$$

La norma ora introdotta è moltiplicativa, cioè

$$N((a + b\sqrt{d})(c + e\sqrt{d})) = N(a + b\sqrt{d})N(c + e\sqrt{d})$$

Sussiste il seguente fatto.

Proposizione 2.3.1. *Gli elementi invertibili di* $\mathbb{Z}\left[\sqrt{d}\right]$, *d non-quadrato in* $\mathbb{Z}$, *sono tutti e soli gli elementi* $a+b\sqrt{d}$ *tali che* $N(a+b\sqrt{d}) = a^2-db^2 = \pm 1$, *cioè*

$$U(\mathbb{Z}\left[\sqrt{d}\right]) = \left\{a + b\sqrt{d} | N(a + b\sqrt{d}) = a^2 - db^2 = \pm 1\right\}.$$

Dimostrazione. Sia $x = a + b\sqrt{d}$ un elemento invertibile (e $x' = a - b\sqrt{d}$ il suo coniugato), detto x^{-1} il suo inverso, sarà

$$1 = N(1) = N(xx^{-1}) = N(x)N(x^{-1})$$

da cui $N(x) = \pm 1$. Viceversa, se x è tale che $N(x) = \pm 1$, allora

$$\pm 1 = N(x) = xx'$$

da cui x è invertibile (se $N(x) = 1$, l'inverso di x è x', se $N(x) = -1$, l'inverso di x è $-x'$). □

Ma dire che la norma di un elemento $a+b\sqrt{d}$ vale ± 1, equivale a dire che $a^2 - db^2 = \pm 1$.

Quindi $a + b\sqrt{d}$ è invertibile se e solo se (a, b) risolve l'equazione (2.1.1).

Se esiste una soluzione (a, b) della (2.1.6) allora:

- le potenze dispari di $a + b\sqrt{d}$ sono soluzioni della (2.1.6);
- le potenze pari sono invece soluzioni della (2.0.1).

Se invece (a, b) è una soluzione della (2.0.1):

- tutte le potenze di $a + b\sqrt{d}$ sono soluzioni della (2.0.1).

Ciò premesso, passiamo allo studio degli elementi invertibili di $\mathbb{Z}\left[\sqrt{d}\right]$, $d \in \mathbb{Z}$ e non quadrato. Si tratta di risolvere le due equazioni

$$a^2 - db^2 = 1 \qquad \text{e} \qquad a^2 - db^2 = -1$$

cercando soluzioni intere.

1. $d = -1$ (caso degli interi di Gauss): si tratta di risolvere in $\mathbb{Z}$ le equazioni

   $$a^2 + b^2 = 1 \qquad \text{e} \qquad a^2 + b^2 = -1.$$

   Soltanto la prima equazione ammette soluzioni: $a = \pm 1$, $b = 0$, o $a = 0$ e $b = \pm 1$. Quindi gli elementi invertibili di $\mathbb{Z}[i]$ sono ± 1 e $\pm i$.

2. $d < -1$ (caso, ad esempio, di $\mathbb{Z}[-3]$): si ha $a^2 - db^2 = a^2 + |d|b^2$ e quindi la $a^2 - db^2 = -1$ non ammette soluzioni. La $a^2 - db^2 = 1$ è soddisfatta solo se $b = 0$ e $a = \pm 1$. Quindi gli unici elementi invertibili sono ± 1.

3. $d > 0$: questo è il caso più complesso. Osserviamo innanzitutto che se z è un'unità, anche $-z$ e $\frac{1}{z}$ lo sono. Inoltre, se $|z| > 1$ risulta $|\frac{1}{z}| < 1$. Quindi le unità diverse da ± 1 di $\mathbb{Z}\left[\sqrt{d}\right]$ sono del tipo $\pm z$,$\pm z^{-1}$. Se esiste una unità di $\mathbb{Z}\left[\sqrt{d}\right]$ diversa da ± 1, possiamo supporre che $z > 1$. Esaminiamo l'equazione

$$a^2 - db^2 = 1 \qquad d > 0,$$

essa ammette sempre una soluzione non banale, ossia diversa da ± 1. Sia allora u la più piccola unità maggiore di 1: una tale soluzione si può ottenere ad esempio ponendo nell'equazione $a^2 - db^2 = 1$ via via $b = 1, 2, \ldots$ finchè $1 + db^2$ non diventi un quadrato perfetto.
Ovviamente tutte le potenze di u^n, $n \in \mathbb{N}$ sono ancora unità maggiori di 1.

Sussiste il seguente teorema.

Teorema 2.3.1. *Sia u il più piccolo elemento invertibile maggiore di 1 dell'anello $\mathbb{Z}\left[\sqrt{d}\right]$, $d > 0$. Allora tutti gli elementi invertibili maggiori di 1 di $\mathbb{Z}\left[\sqrt{d}\right]$ sono del tipo u^n, $n \in \mathbb{N}$.*

Dimostrazione. Supponiamo per assurdo che esista una unità $x > 1$ che non sia una potenza di u. Allora x si troverà tra due potenze successive di u, ossia esisterà un intero positivo m tale che

$$u^m < x < u^{m+1}.$$

Moltiplicando tali disuguaglianze per u^{-m} si ottiene

$$1 < u^{-m}x < u.$$

Avremo trovato una unità, $u^{-m}x$, strettamente contenuta tra 1 e u, il che contraddice la scelta di u. Quindi ogni unità maggiore di 1 è una potenza ad esponente intero positivo di u. □

Bibliografia

[1] Cristiano Teodoro. *Sulla risoluzione delle equazioni di Pell.* http://matematicamente.it/approfondimenti/Equazioni.pdf

[2] Kiran S.Kedlaya. *Solving constrained Pell equations.* http://math:princeton.edu/kkledlaya.pdf

[3] Harold Davenport. *Aritmetica superiore. Un'introduzione alla teoria dei numeri.* Zanichelli Editore S.p.A., Bologna (sesta edizione 1994).

[4] K.F. Ireland-M.I. Rosen. *A classical introduction to modern number theory.* Springer, (1982).

[5] H.E. Rose. *A course in number theory.* Oxford Science Publ., (1988).

[6] Havil, J. Gamma. *Exploring Euler's Constant.* Princeton,NJ: Princeton University Press,2003.

[7] D'Angelo,J.P. and West, D.B.*Mathematical Thinking:Problem-Solving and Proofs.* Uper Saddle River,NJ: Prentice-Hall (seconda edizione,2000).

[8] C.D.Olds. *Continued fractions.* random Hause and The L.W. Singer Company, (second Printing, 1963).

[9] G.H. Hardy and E.M. Wright. *An introdution to the theory of numbers.* The Clarendon Press, Oxford University Press, New York, (fifth edition, 1979).

[10] T. Nagel. *Introduction to Number Theory.* New York: Wiley, 1951.