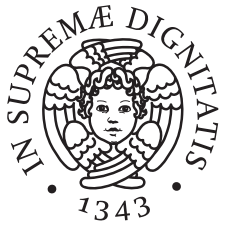

Università degli Studi di Pisa
Facoltà di Scienze Matematiche, Fisiche e Naturali
Corso di Laurea in Matematica

Tesi di Laurea Triennale

Equazione di Pell in Interi e Polinomi

22 Settembre 2017

Candidato
Dario Balboni

Relatore
Prof. **Roberto Dvornicich**
Università di Pisa

Anno Accademico 2016/2017

Indice

Capitolo 1

Introduzione

1.1 Overview della Tesi

Ci poniamo in questa sede l'obiettivo di fornire una breve introduzione all'equazione di Pell, trattandone differenze e similitudini delle due varianti principali: sugli interi e sui polinomi, con particolare riferimento alla risoluzione esplicita di essa e qualche osservazione relativa alla specificazione delle soluzioni minime. Più nel dettaglio, nel capitolo 1 viene data una veloce introduzione storica all'equazione di Pell, mentre nel capitolo 2 vengono descritti i motivi per i quali ancora oggi merita di essere studiata.

Nel capitolo 3 vengono osservate alcune proprietà della struttura delle soluzioni dell'equazione ed alcune differenze dell'equazione polinomiale dalla sua controparte intera, come l'influenza dei fattori quadratici, ed un criterio di non risolubilità basato sul teorema di Mason-Stothers.

Nel capitolo 4 viene introdotto lo strumento delle frazioni continue, molto importanti in tutta la teoria dell'approssimazione diofantea. Oltre alle principali formule aritmetiche, vengono esposti anche i teoremi riguardanti la bontà delle approssimazioni dovuti alle convergenti delle frazioni continue e le applicazioni di questi ultimi alla risoluzione dell'equazione di Pell. Per fare ciò viene definita la radice quadrata di un polinomio nel campo delle serie di Laurent e viene dato un algoritmo per permettere il calcolo esplicito dello sviluppo in frazioni continue della radice quadrata.

Successivamente, sempre nello stesso capitolo, si espone un famoso teorema dovuto a Lagrange che lega la periodicità dello sviluppo in frazione continua all'essere irrazionali quadratici. Viene inoltre fornito il principale algoritmo risolutivo che utilizza le frazioni continue.

Nel capitolo 5 viene fornita una condizione necessaria e sufficiente per verificare l'esistenza di soluzioni con norma negativa e viene analizzato il metodo detto "dei sistemi di riduzione", che consente di risolvere anche l'equazione di Pell generica. Nello stesso capitolo viene delineato il collegamento tra le soluzioni polinomiali ed i divisori di torsione sulla varietà jacobiana di particolari curve iperellittiche associate all'equazione stessa.

Infine, nel capitolo 6 si discutono alcuni risultati sulla valutazione di soluzioni polinomiali minime in $\mathbb{Z}[t]$ per ottenere soluzioni minime in $\mathbb{Z}$.

1.2 Cos'è l'Equazione di Pell

Fissato un anello R, consideriamo $d \in R$. La seguente equazione viene detta "equazione di Pell":

$$y^2 - dx^2 = c \tag{1.1}$$

Viene chiesto, per $c \in R$ fissato (solitamente 1), di discutere l'esistenza di soluzioni $x, y \in R$ al variare di $d \in R$.

In questa tesi discuteremo principalmente i casi in cui $R = \mathbb{Z}$ oppure $R = K[t]$ con K campo, seppur alcuni teoremi abbiano una validità più generale.

Le motivazioni matematiche per le quali essa viene ancora studiata verranno fornite nella prossima sezione, mentre in questa vorrei dare un'introduzione storica al problema, tratta da [4, Capitolo 2].

1.2.1 Archimede ed il problema del Bestiame

È interessante notare come già Archimede pose, in una lettera indirizzata ad Eratostene di Cirene, un problema - successivamente nominato "Il problema del Bestiame" - nel quale veniva chiesto di calcolare il numero di capi di bestiame presenti nell'isola della Sicilia, sapendo che esso era soggetto ad alcune condizioni. Riscritto in linguaggio matematico moderno, esso chiede di risolvere sugli interi un sistema di equazioni in otto incognite. Sette equazioni sono lineari omogenee a coefficienti razionali e per queste, attraverso semplice algebra lineare, si riesce a trovare una forma esplicita che parametrizza tutte le soluzioni; le ultime due sono invece condizioni quadratiche: si richiede infatti che la somma di due delle otto variabili sia un quadrato perfetto e che la somma di altre due sia un numero triangolare.

Con poche manipolazioni algebriche si può ricondurre il problema al calcolo delle soluzioni di una equazione di Pell sugli interi per $d = 410286423278424$, numero enorme per le capacità computazionali dell'epoca.

1.2.2 Contributo dei Greci

La prima menzione esplicita di una equazione di Pell pare essere occorsa nel lavoro di Theone di Smyrna, che diede delle formule per ricorrenza per generare soluzioni dell'equazione di Pell sugli interi per $d = 2$. Proclo (filosofo neoplatonico) le ridusse successivamente all'identità

$$(2a + b)^2 - 2(a + b)^2 = -(a^2 - 2b^2)$$

da cui si può derivare una dimostrazione del fatto che

$$s_{n+1} = s_n + d_n, \qquad d_{n+1} = 2s_n + d_n$$

sono soluzioni dell'equazione di Pell con $d = 2$ (per n pari) se partiamo da $s_1 = d_1 = 1$, seppur tale dimostrazione non compaia nel lavoro di Proclo.

Straordinariamente pare che i Pitagorici usassero i valori di $\frac{d_n}{s_n}$ come mezzo per produrre approssimazioni sempre migliori di $\sqrt{2}$. Come vedremo nel Capitolo 4 esse tendono proprio a $\sqrt{2}$, ma i Pitagorici non ne possedevano alcuna dimostrazione.

1.2.3 I matematici Indiani

Ci fu grande interesse in generale per le equazioni diofantee tra i matematici indiani. Ad esempio, già nel quinto secolo dopo cristo, Aryabhata I aveva sviluppato un metodo per risolvere l'equazione diofantea lineare

$$ax - by = c$$

per $x, y \in \mathbb{Z}$.

Nel 628 Brahmagupta fu il primo a scoprire la cosiddetta identità di moltiplicazione: se $A^2 - DB^2 = Q$ e $P^2 - DR^2 = S$, allora si ha $(AP + DBR)^2 - D(AR + BP)^2 = QS$, che permette, note due soluzioni della Pell, di ottenerne una terza.

Ma il più grosso risultato dei matematici Indiani relativamente all'equazione di Pell fu lo sviluppo del metodo ciclico per risolverla. La tecnica, descritta da Bhaskara II nel 1150, consiste nel ridurre la soluzione della Pell per c generico alla risoluzione di una equazione di Pell con lo stesso d ma con $c = 1, 2, 4$, per le quali già Bramagupta aveva dato un metodo risolutivo. Gli Indiani non dimostrarono il loro risultato, ma si accontentarono di utilizzarlo per risolvere l'equazione di Pell per $d = 61, 67, 97, 103$, e fu solo nel 1930 che Ayyangar diede una dimostrazione della validità del metodo ciclico.

1.2.4 Fermat ed i suoi successori

Nel 1657 Fermat, in una lettera a Frénicle ed ai matematici in generale, chiese una dimostrazione dell'esistenza di infinite soluzioni per $d \in \mathbb{N}$ non quadrato perfetto. Successivamente chiese se esistesse un modo per poter trovare le soluzioni dell'equazione.

Brouckner trovò un metodo alquanto ingegnoso, che fu poi modificato ed esteso da Eulero che si accorse dell'importanza delle frazioni continue nel dare un algoritmo efficiente per il calcolo delle soluzioni, ma non riuscì comunque a dimostrare che in questo modo si poteva ottenere una soluzioni per tutti i d non quadrati, risultato che fu invece successivamente ottenuto da Lagrange.

1.2.5 Il contributo di Pell

Per quanto possa sembrare strano, Pell non diede alcun contributo fondamentale all'equazione che porta il suo nome. Il motivo per cui viene oggigiorno chiamata così è dovuto ad un errore di Eulero, che attribuì il metodo risolutivo di Brouckner a John Pell.

Capitolo 2

Importanza dell'Equazione di Pell

Vediamo ora i motivi per cui l'equazione di Pell viene studiata anche oggi: alcuni problemi abbastanza naturali in teoria dei numeri si riescono a ridurre ad essa, consentendo così di risolverli; è inoltre legata al class number dei campi quadratici reali dalla class number formula e consente di descrivere le unità negli stessi. Nella versione polinomiale essa è legata anche a questioni di esistenza di primitive di integrali di funzioni algebriche esprimibili in termini di funzioni elementari.

2.1 Riduzione delle equazioni di Secondo Grado

Osservazione 1. Consideriamo una generica equazione di secondo grado

$$ax^2 + bxy + cy^2 + dx + ey + f = 0 \qquad a, b, c, d, e, f \in R \tag{2.1}$$

Attraverso semplici cambi di variabili per completamento dei quadrati, si può portare l'equazione di sopra nella forma dell'equazione 1.1 se $2 \nmid \operatorname{char} R$ e $\Delta = b^2 - 4ac \neq 0$ e $a \neq 0$.

Dimostrazione. Moltiplicando l'equazione 2.1 per Δ^2 ed effettuando il cambio di variabili $u = \Delta x - H$, $v = \Delta y - K$ con $H, K \in R$ da determinare, si può verificare che per $H = -eb + 2cd$ e $K = -bd + 2ae$ si riescono ad annullare entrambi i termini lineari, ottenendo quindi l'equazione

$$au^2 + buv + cv^2 + \Delta^2 f = 0 \tag{2.2}$$

A questo punto moltiplicando tutta l'equazione per $4a$ e ponendo $w = 2au + bv$ si ottiene

$$w^2 - \Delta v^2 = -\Delta^2 f \tag{2.3}$$

che riconosciamo essere un'equazione di Pell.

Risolvendo quest'ultima e ricordandosi dei cambi di variabili fatti (uno dei quali prevede di risolvere una diofantea lineare) si possono ritrovare tutte le soluzioni all'equazione originaria. □

2.2 Unità dei campi quadratici e class number

Data un'equazione di Pell per $d \in R$, possiamo considerare $R[\sqrt{d}] = \frac{R[x]}{(x^2-d)}$, anello quadratico su R. Ogni elemento $\alpha \in R[\sqrt{d}]$ si scrive in modo unico come $\alpha = p + \sqrt{d}q$, con $p, q \in R$.

Un elemento $\alpha \in R[\sqrt{d}]$ è un'unità, ovvero $\exists \beta \in R[\sqrt{d}]$ tale che $\alpha\beta = 1$ se e solo se la sua "norma" $\mathcal{N}(p + \sqrt{d}q) = (p + \sqrt{d}q)(p - \sqrt{d}q) = p^2 - dq^2$ (che osserviamo essere moltiplicativa) è un invertibile di R.

Infatti $1 = \mathcal{N}(1) = \mathcal{N}(\alpha\beta) = \mathcal{N}(\alpha)\mathcal{N}(\beta)$. Inoltre se $\mathcal{N}(\alpha) = e$ è invertibile, possiamo considerare $\beta = e^{-1}(p - \sqrt{d}q)$ e quindi si ha $\alpha\beta = e^{-1}\mathcal{N}(\alpha) = 1$.

In questo senso, le soluzioni all'equazione di Pell con c invertibile rappresentano le unità dell'anello $R[\sqrt{d}]$.

2.2.1 Connessione con il class number

Nel caso in cui $R = \mathbb{Q}$ e $0 < d \in \mathbb{N}$, $R[\sqrt{d}]$ è un campo quadratico reale e la Class Number Formula per i campi quadratici dice che

$$h(d) = \frac{\sqrt{d}}{\ln \varepsilon} L(1, \chi) \tag{2.4}$$

dove $h(d)$ è il class number di $\mathbb{Q}(\sqrt{d})$, ε viene detta unità fondamentale, e $L(s, \chi)$ è la L-serie di Dirichlet del campo. Dirichlet mostrò anche che la L-serie per questi campi si può scrivere in forma finita, ovvero supponendo che χ sia un carattere primitivo con conduttore primo q si ha

$$L(1, \chi) = \begin{cases} -\frac{\pi}{q^{\frac{3}{2}}} \sum_{m=1}^{q-1} m\left(\frac{m}{q}\right) & \text{se } q \equiv 3 \pmod 4 \\ -\frac{1}{q^{\frac{1}{2}}} \sum_{m=1}^{q-1} \left(\frac{m}{q}\right) \ln(2) \sin\left(\frac{m\pi}{q}\right) & \text{se } q \equiv 1 \pmod 4 \end{cases}$$

Pertanto, visto che $h(d)$ è un numero naturale ed $L(1, \chi)$ si può calcolare dalle formule precedenti con precisione arbitraria, se fosse noto $\ln \varepsilon$ con sufficiente precisione, si potrebbe determinare il class number di $\mathbb{Q}(\sqrt{d})$.

L'unità fondamentale ε si può caratterizzare come

$$\varepsilon = \frac{a + b\sqrt{d}}{2}$$

dove (a, b) è la soluzione più piccola a

$$x^2 - dy^2 = \pm 4$$

negli interi positivi. Questo legame con l'equazione di Pell mostra quanto importante sia riuscire a risolverla.

2.3 Primitive esatte di funzioni algebriche

Nel caso in cui $R = K[x]$ con K campo, già Abel [1] notò, qualche secolo fa, che se $p(x)$ e $q(x)$ risolvono l'equazione di Pell con $d(x)$, allora vale che

$$\int \frac{p'(x)\mathrm{d}x}{q(x)\sqrt{d(x)}} = \log\left(p(x) + q(x)\sqrt{d(x)}\right) \tag{2.5}$$

, espressione che generalizza il ben noto integrale

$$\int \frac{\mathrm{d}x}{\sqrt{x^2+2bx+c}} = \log\left(x+b+\sqrt{x^2+2bx+c}\right)$$

Ciò è piuttosto sorprendente, poiché in generale ci si aspetterebbe di ottenere funzioni ellittiche inverse o peggio, piuttosto che un logaritmo di una funzione algebrica.

Capitolo 3

Primi risultati sull'Equazione

In tutto questo capitolo, e nei seguenti per equazione di Pell intenderemo l'equazione 1.1 con $c = 1$, se non diversamente specificato. Per risolvere il caso $c \neq 1$ risulta comodo aver risolto il caso per $c = 1$ e pertanto verrà trattato in seguito utilizzando il metodo dei sistemi di riduzione.

Inoltre supporremo sempre che char $K \neq 2$: in questo caso infatti si hanno svariati problemi che non siamo interessati a trattare.

3.1 Preliminari e Prime Osservazioni

I casi di nostro interesse per la risoluzione dell'equazione di Pell saranno $R = \mathbb{Z}$ e $R = K[t]$ con K campo, anche se molti risultati che otterremo saranno più generali.

3.1.1 Il gruppo delle soluzioni

Lemma 1: Sia $S \subseteq R$ un insieme moltiplicativamente chiuso[1] di R. A D fissato definiamo l'insieme delle soluzioni a valori in S: $\mathrm{Sol}_S = \{(x, y) \in R \mid x^2 - Dy^2 \in S\}$.

L'insieme Sol_S ha una struttura di monoide abeliano dove l'elemento neutro è $(1, 0)$.

Dimostrazione. Dati $(x_1, y_1), (x_2, y_2) \in \mathrm{Sol}_S$ definiamo, utilizzando un'identità già scoperta da Brahmagupta,

$$(x_1, y_1) \star (x_2, y_2) = (x_1x_2 + Dy_1y_2, x_1y_2 + x_2y_1) \tag{3.1}$$

infatti si ha $(x_1x_2 + Dy_1y_2)^2 - D(x_1y_2 + x_2y_1)^2 = (x_1^2 - Dy_1^2)(x_2^2 - Dy_2^2)$ e la tesi segue poiché S è moltiplicativamente chiuso. □

Osservazione 2. Se $S = \{1\}$ oppure $S = R^*$, Sol_S ha in realtà una struttura di gruppo con inverso dato da

$$(x_1, y_1)^{-1} = (sx_1, -sy_1)$$

dove $s = (x_1^2 - Dy_1^2)^{-1}$

[1] Ovvero vale $\forall a, b \in S \quad ab \in S$, $1 \in S$, $0 \notin S$

3.1.2 Caso intero

Sugli interi consideriamo $\mathbb{Z} \subseteq \mathbb{R}$, come grado la funzione modulo $|\cdot| : \mathbb{Z} \to \mathbb{N}$ e come funzione di approssimazione $\lfloor \cdot \rfloor : \mathbb{R} \to \mathbb{Z}$ la funzione "intero più vicino di modulo minore", ovvero

$$\lfloor x \rfloor = n \in \mathbb{Z} \text{ tale che } \forall m \neq n \in \mathbb{Z} \quad |x - m| < |x - n| \implies |m| > |n| \tag{3.2}$$

3.1.3 Caso polinomiale

Sui polinomi consideriamo invece $K[t] \subseteq K((t^{-1}))$ dove $K((t^{-1}))$ è il campo delle serie di Laurent, ovvero delle scritture formali del tipo $\sum_{h=-k}^{\infty} f_h t^{-h}$ (notiamo che hanno solo un numero finito di termini positivi e possono avere infiniti termini negativi) dotate del prodotto alla Cauchy[2]. Definiamo inoltre la funzione modulo come "esponenziale del grado del coefficiente massimo", ovvero se $F = \sum_{h=-k}^{\infty} f_h t^{-h}$

$$|F| = e^{\min\{h | f_h \neq 0\}} \tag{3.3}$$

e come funzione di approssimazione quella che restituisce le parti polinomiali della serie di Laurent considerata:

$$\lfloor F \rfloor = \sum_{h=-k}^{0} f_h t^{-h} \tag{3.4}$$

Notiamo che la nozione di grado usuale per i polinomi viene a coincidere con la nozione appena fornita per le serie di Laurent (a meno di un logaritmo, ovvero $\deg P = \log |P|$). Il motivo di questa strana scelta potrà essere apprezzato in seguito, in quanto consentirà di enunciare in maniera uniforme i teoremi in comune tra interi e polinomi.

Esempio 1: Considerando $F = 3T^5 + 4 - 7T^{-8}$ si ha che $|F| = e^5$ e che $\lfloor F \rfloor = 3T^5 + 4$.
Se si considera invece $F = 5T^{-2} + 3T^{-7}$ si ha $|F| = e^{-2}$ e $\lfloor F \rfloor = 0$.

Condizioni necessarie per la risolubilità

Notiamo subito che, detto $a = \deg d \geq 1$ nell' equazione di Pell 1.1, $n = \deg x$, $m = \deg y$, affinché avvenga cancellazione, deve essere $2n = a + 2m$ e quindi $\deg d$ deve essere pari.

Passando all'uguaglianza tra i coefficienti direttivi dei tre polinomi (x^2, dy^2 e 1) si vede che il coefficiente direttivo di d deve essere un quadrato in K, affinché l'equazione ammetta soluzioni in $K[t]$.

3.1.4 d quadrato perfetto

Nel caso in cui R sia un dominio, $S = R^*$, supponiamo che $d = r^2$ per un qualche $r \in R$. Allora presa una soluzione $(x, y) \in \mathrm{Sol}_S$ si ha che la 1.1 diventa

$$(x + ry)(x - ry) = e$$

[2]Il prodotto alla Cauchy si definisce come $(\sum_i a_i t^i)(\sum_j b_j t^j) = \sum_k c_k t^k$ con $c_k = \sum_{i+j=k} a_i b_j$.

ovvero si ha $x + ry = ea$, $x - ry = a^{-1}$ con $e, a \in R^*$ da cui, facendone le semisomme e semidifferenze si ottiene

$$2x = ea + a^{-1}, \qquad 2ry = ea - a^{-1}$$

E quindi otteniamo tutte le soluzioni della 1.1 su R come $(\frac{ea+a^{-1}}{2}, \frac{ea-a^{-1}}{2r})$, al variare di $a \in R^*$ (anche se non tutte le coppie scritte in quella forma sono soluzioni su R).

Osservazione 3. Se $R = \mathbb{Z}$ oppure $R = K[x]$ le uniche soluzioni dell'equazione di Pell con $c \in R^*$ e d quadrato perfetto sono banali.

Dimostrazione. Nel caso di $\mathbb{Z}$, utilizzando il ragionamento e le notazioni di sopra, e ricordando che gli invertibili di $\mathbb{Z}$ sono solo ± 1, otteniamo che

- $x + ry = -1$ e $x - ry = -1$, nel qual caso $x = -1$ e quindi necessariamente $y = 0$
- $x - ry = 1$ e $x + ry = 1$, nel qual caso $x = 1$ e quindi $y = 0$

Nel caso di $K[x]$ invece, sempre utilizzando il ragionamento già svolto e ricordando che gli invertibili sono i polinomi costanti (ovvero quelli il cui grado è nullo) si ha che $|2x| = \left|ea + a^{-1}\right| = 1$ e quindi x è un polinomio costante. Da cui si ha che si ha che $y = 0$ oppure r è costante e quindi anche $d = r^2$ risulta costante: assurdo. □

3.2 Alcuni teoremi sulle soluzioni

3.2.1 Struttura delle soluzioni

Il gruppo abeliano delle soluzioni corrispondenti a $c = 1$, sia nel caso $R = \mathbb{Z}$ che nel caso $R = K[x]$, è isomorfo a $\frac{\mathbb{Z}}{2\mathbb{Z}} \times \frac{\mathbb{Z}}{2\mathbb{Z}} \times \mathbb{Z}$:

Osservazione 4. Se (x, y) e (u, v) sono due soluzioni di 1.1 con $c = 1$, così lo sono anche

$$(xu + dyv, xv + uy) \qquad \text{e} \qquad (xu - dyv, yu - vx)$$

ed inoltre si può notare che

$$(xv + uy)(uy - xv) = y^2 - v^2$$

Teorema 2: Sia $d \in \mathbb{Z}$, $d > 0$ e non quadrato perfetto. Allora se l'equazione ha soluzione, esiste una soluzione (x, y) non banale a

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

con $x, y \in \mathbb{Z}$, tale che tutte le soluzioni non banali sono della forma $(\pm p, \pm q)$ dove

$$p + \sqrt{d}q = (x + \sqrt{d}y)^n$$

per un qualche $n \in \mathbb{N}$.

Una tale soluzione (x, y) viene detta minima. Esiste un'unica soluzione minima a meno di cambi di segno.

Dimostrazione. Consideriamo l'insieme delle soluzioni (u, v) dell'equazione con $u, v > 0$ (non vuoto per ipotesi) e diciamo minima la soluzione (x, y) con seconda componente minima tra tutte le soluzioni.

Notiamo che, siccome $x^2 = 1 + dy^2$, la richiesta che la seconda componente sia minima ci dà automaticamente che anche la prima componente della coppia è minima tra tutte le prime componenti delle soluzioni e che la soluzione con seconda componente minima è unica, visto che la prima componente è univocamente determinata dalla scelta della seconda. Da ciò segue quindi che anche $x + \sqrt{d}y$ è minimo tra tutte le espressioni così ottenute al variare delle soluzioni, essendo $\sqrt{d} > 0$.

Mostriamo la tesi per induzione forte sulla seconda componente della soluzione. Presa una qualunque soluzione (p, q) con $p, q > 0$ si avrà allora che $\exists! m \in \mathbb{N}$ tale che

$$(x + \sqrt{d}y)^m \leq (p + \sqrt{d}q) < (x + \sqrt{d}y)^{m+1}$$

Utilizzando l'osservazione precedente con la soluzione minima e con la soluzione (p, q), si ha che almeno uno tra $xq + yp$ e $xq - yp$ è minore stretto di q:

- Nel primo caso per ipotesi induttiva abbiamo che

  $$(xp + dyq) + \sqrt{d}(xq + yp) = (x + \sqrt{d}y)^n$$

  per qualche $n \in \mathbb{N}$, ovvero

  $$p + \sqrt{d}q = (x + \sqrt{d}y)^{n-1}$$

- Nel secondo caso invece si ha

  $$(xp - dyq) + \sqrt{d}(yp - xq) = (x + \sqrt{d}y)^n$$

  per qualche $n \in \mathbb{N}$, e quindi

  $$p - \sqrt{d}q = (x + \sqrt{d}y)^n$$

  ovvero $p + \sqrt{d}q = (x - \sqrt{d}y)^{-n}$

□

Teorema 3: Sia $d \in K[x]$ un polinomio tale che l'equazione 1.1 con $c = 1$ abbia soluzioni non banali in $K[x]$. Allora esiste una soluzione minimale (x, y) tale che ogni altra soluzione non banale è della forma $(\pm p, \pm q)$ dove

$$p + \sqrt{d}q = (x + \sqrt{d}y)^n$$

per qualche $n \in \mathbb{N}$. Una tale soluzione viene detta minima; esiste un'unica soluzione minima a meno di cambi di segno.

Inoltre se $d \in T[x]$, con T un sottoanello di K, è tale che l'equazione di Pell ha una soluzione non banale in $T[x]$, esiste un intero positivo m tale che ogni altra soluzione non banale in $T[x]$ dell'equazione è della forma $(\pm p, \pm q)$ dove

$$p + \sqrt{d}q = (x + \sqrt{d}y)^{mn}$$

per qualche $n \in \mathbb{N}$.

Dimostrazione. Diciamo minima una soluzione in cui il grado del polinomio nella seconda componente sia minimo tra tutte le soluzioni e notiamo, come sopra, che questo garantisce che anche la prima componente abbia grado minimo tra tutte le soluzioni.

Non ci garantisce invece l'unicità, visto che ci sono più polinomi dello stesso grado fissato: supponiamo allora di avere due soluzioni minime (x, y) e (X, Y) e sia $s = \deg y = \deg Y$. Utilizzando l'osservazione precedente si ottiene

$$\deg(xY + Xy) + \deg(xY - Xy) \leq 2s$$

- Se entrambi i gradi nel membro di sinistra fossero uguali ad s, si avrebbe

  $$\deg(2Yx) = \deg(Xy + Yx + Yx - Xy) \leq s$$

  da cui seguirebbe che $\deg x = 0$, una contraddizione.

- Quindi si ha, ad esempio, che $\deg(Yx - Xy) < s$ (l'altro caso può essere trattato analogamente). Allora la soluzione $(Xx - dYy, Yx - Xy)$ deve essere la soluzione banale, ovvero $Yx = Xy$ e $Xx = \pm 1 + dYy$.

  Poniamo allora λ tale che $X = \lambda x$ e $Y = \lambda y$ (che esiste per la prima delle due equazioncine trovate) ed abbiamo $\lambda x^2 = \pm 1 + \lambda d y^2$ e quindi $\lambda = \lambda(x^2 - dy^2) = \pm 1$, da cui si ha $(X, Y) = (x, y)$ oppure $(X, Y) = (-x, -y)$ e questo dimostra la prima parte del teorema.

Per la seconda parte assumiamo per induzione forte, come già fatto nel caso degli interi, che tutte le soluzioni dell'equazione di Pell sui polinomi con seconda componente di grado $< l$ siano della forma desiderata.

Consideriamo allora una soluzione (p, q) con $\deg q = l$. Con lo stesso ragionamento di prima si hanno i due casi $\deg(py + qx) < l$ oppure $\deg(qx - py) < l$.

E la dimostrazione si conclude come nel caso dei numeri interi.

Per la parte relativa alle soluzioni su un sottoanello notiamo che ogni soluzione in $T[x]$ deve essere anche una soluzione in $K[x]$ e quindi si scriverà come una certa potenza naturale della soluzione minima su $K[x]$.

Consideriamo allora la minima potenza m per la quale $(x + \sqrt{d}y)^m$ ci da l'espressione $p + \sqrt{d}q$ con $p, q \in T[x]$. Ne segue che $(x + \sqrt{d}y)^{mn}$ sono soluzioni dell'equazione di Pell per ogni $n \in \mathbb{N}$.

Se ci fosse una soluzione in $T[x]$ diversa da queste, diciamo ad esempio $(x + \sqrt{d}y)^{mn+r}$ per qualche $0 < r < m$, allora si avrebbe

$$(x + \sqrt{d}y)^r = (x + \sqrt{d}y)^{mn+r}(x - \sqrt{d}y)^{mn}$$

che darebbe una soluzione in $T[x]$ in contraddizione con la minimalità di m. □

3.2.2 Particolarità dell'equazione polinomiale

L'equazione polinomiale, pur avendo molti tratti in comune con l'equazione sugli interi, ha anche diverse caratteristiche che la differenziano: in particolare il fatto che essa non sia sempre risolubile (come invece vedremo accadrà per l'equazione intera) e che i fattori quadratici giochino una certa influenza.

Soluzioni per $\deg d = 2$

Nel caso in cui il grado del polinomio d sia 2 e d non sia un quadrato perfetto, siamo capaci di dare esplicitamente la forma delle soluzioni: se $d = c^2(t-\alpha)(t-\beta) \in K[t]$ con $c \neq 0$ e $\alpha \neq \beta$ si ha

$$P = \frac{2t - (\alpha+\beta)}{\alpha - \beta}, \qquad Q = \frac{2}{c(\alpha-\beta)}$$

è una soluzione e avendo Q grado zero, è la minima.

Metodo risolutivo con stime a priori sul grado

Si potrebbe pensare che, se avessimo stime a priori sul grado delle soluzioni (diciamo k) in funzione del grado del polinomio d, potremmo scrivere - a d fissato - l'equazione di Pell ponendo $x = \sum_{i=0}^{k} a_i t^i$ e $y = \sum_{i=0}^{k} b_i t^i$ (dove i coefficienti a_i, b_i sono pensati come incognite), aprire l'equazione risultante ed imporre l'uguaglianza tra i coefficienti dei polinomi, ottenendo così un sistema di equazioni algebriche in $2k$ incognite su K, per le quali esistono metodi algoritmici noti per sapere se hanno soluzione oppure no, se K è algebricamente chiuso.

Purtroppo non si riescono a dare stime a priori sul grado; ad esempio già per $\deg d = 4$ il grado delle soluzioni minime può essere arbitrariamente alto. Si veda a questo proposito il controesempio alla fine dell'articolo [3].

Particolarità dei fattori quadratici

I fattori quadratici, che nel caso intero come vedremo non influiscono sull'esistenza di soluzioni, nel caso polinomiale sono più rilevanti. Consideriamo la seguente affermazione, che saremo in grado di dimostrare tra poco:

Fatto 1: Data una coppia di polinomi fissati $R(t), S(t)$, si può considerare $d = R(t)S(t)^{2k}$. Per ogni coppia di polinomi fissati, esiste $k \in \mathbb{N}$ tale che l'equazione di pell per d non ammetta soluzioni non banali.

Per un esempio più esplicito si può considerare $d = (t-\lambda)^4(t^2-1)$ per $\lambda \in \mathbb{Q}$.

Condizione necessaria per la risolubilità

Teorema 4 (Mason-Stothers)**:** Se A, B, C sono polinomi coprimi su $K[t]$, tali che $A + B = C$, ed almeno una derivata A', B', C' non è completamente nulla allora

$$\max\{\deg A, \deg B, \deg C\} < n(ABC) \tag{3.5}$$

dove $n(P)$ denota il numero di zeri distinti di P in una chiusura algebrica.

La dimostrazione è tratta da [11], benché originariamente il risultato fosse dovuto indipendentemente a Mason [6] e Stothers [12]. Per arrivarci ci servirà un lemma preliminare:

Lemma 5: Sia f un polinomio non nullo in $K[x]$, con K algebricamente chiuso. Allora

$$\deg f \leq \deg \mathrm{MCD}(f, f') + n(f)$$

dove $n(f)$ è il numero di zeri distinti di f e f' è la derivata formale di f.

Dimostrazione. Siano $\alpha_1, \ldots, \alpha_m$ le radici di f con molteplicità $a_1, \ldots, a_m$ in modo che $f = c(x-\alpha_1)^{a_1} \cdot \ldots \cdot (x-\alpha_m)^{a_m}$. Allora, per la regola del prodotto si ha

$$f' = ca_1(x-\alpha_1)^{a_1-1}(x-\alpha_2)^{a_2} \cdot \ldots \cdot (x-\alpha_m)^{a_m} + c(x-\alpha_1)^{a_1}\frac{\mathrm{d}}{\mathrm{d}x}((x-\alpha_2)^{a_2} \cdot \ldots \cdot (x-\alpha_m)^{a_m})$$

da cui si ottiene $(x-\alpha_1)^{a_1-1} \mid \mathrm{MCD}(f, f')$. In maniera simile si ottiene $(x-\alpha_i)^{a_i-1} \mid \mathrm{MCD}(f, f')$ e quindi vediamo che $(x-\alpha_1)^{a_1-1} \cdot \ldots \cdot (x-\alpha_m)^{a_m-1} \mid \mathrm{MCD}(f, f')$. Da ciò segue, siccome f è non nulla, che $\deg(f) - n(f) \leq \deg \mathrm{MCD}(f, f')$. □

Dimostrazione (di Mason-Stothers)**.** Mostreremo, nelle notazioni di sopra, che $\deg(c) \leq n(abc) - 1$: derivando l'uguaglianza nell'ipotesi si ottiene $a' + b' = c'$. Ora, moltiplicando la prima equazione per a', la seconda per a e sottraendo, otteniamo

$$a'b - ab' = a'c - ac'$$

Da questo otteniamo che sia $\mathrm{MCD}(a, a')$ che $\mathrm{MCD}(b, b')$ e pure $\mathrm{MCD}(c, c')$ dividono $a'b - ab'$. Siccome sono tutti e tre relativamente coprimi, otteniamo

$$\mathrm{MCD}(a, a')\mathrm{MCD}(b, b')\mathrm{MCD}(c, c') \mid a'b - ab'$$

Il nostro claim è che $a'b - ab' \neq 0$. Infatti se fosse nullo, avremmo $a \mid a'b$. Siccome a e b sono relativamente coprimi, $a \mid a'$, da qui si ottiene $a' = 0$. Similmente, anche $b' = c' = 0$, contraddicendo l'assunzione. Quindi, il membro di destra è non nullo e si ha

$$\deg \mathrm{MCD}(a, a') + \deg \mathrm{MCD}(b, b') + \deg \mathrm{MCD}(c, c') \leq \deg(a) + \deg(b) - 1$$

Portando ogni cosa a destra ed aggiungendo $\deg(c)$ ad entrambi i membri troviamo che

$$\deg(c) \leq \deg(a) - \deg \mathrm{MCD}(a, a') + \deg(b) - \deg \mathrm{MCD}(b, b') + \deg(c) - \deg \mathrm{MCD}(c, c') - 1$$

Applicando il lemma precedente otteniamo quanto desiderato. □

Lemma 6: Sia K un campo di caratteristica zero. Se vale che

$$n(d) \leq \frac{1}{2}\deg d$$

dove $n(d)$ è il numero di zeri distinti di d in una chiusura algebrica, allora l'equazione di pell polinomiale non ha soluzioni non banali in $K[x]$

Dimostrazione. Applicando il teorema 4 di Mason-Stothers ad una supposta soluzione dell'equazione di Pell con $A = x^2$, $B = -dy^2$ e $C = 1$ otteniamo

$$\deg d + 2\deg y = \deg(dy^2) < n(dx^2y^2)$$

siccome $n(dx^2y^2) = n(dxy) \leq \deg x + n(d) + \deg y$ e $\deg d = 2\deg x - \deg y$, si ha

$$n(d) > \deg x - \deg y = \frac{1}{2}\deg d$$

il che ci dà un assurdo con le ipotesi □

Osservazione 5. Notiamo anche che la stima ottenuta dal lemma è sharp, ovvero non può essere migliorata per tutti i $\deg d$. Infatti si ha l'identità

$$(t^k+1)^2-(t^{2k}+2t^k)\cdot 1^2=1$$

che mostra che esiste una soluzione non banale per $d=t^{2k}+2t^k$, dove $\deg d=2k$ e $n(d)=k+1$

Osservazione 6. La condizione ottenuta non è però sufficiente a garantire l'esistenza di una soluzione: consideriamo $D=(t^2+1)(t-a)^2$ con $a\in\mathbb{R}^+$. L'equazione di pell per questo polinomio non ha soluzioni non banali su $\mathbb{C}[t]$. Ciò non è subito evidente in quanto non è un quadrato perfetto e non si può applicare il lemma appena dimostrato poiché $\deg D=4$ e $n(D)=3$.

Osserviamo però che se abbiamo una soluzione $P,Q\in\mathbb{Q}[t]$ deve valere anche che

$$P(t)^2-(t^2+1)((t-a)Q(t))^2=1$$

, ovvero detto $R(t)=(t-a)Q(t)$, la coppia $(P(t),R(t))$ deve essere soluzione della Pell con $D'(t)=t^2+1$, ovvero in particolare deve valere $R(a)=0$.

Per quanto osservato più in su, le soluzioni di quest'ultima sono tutte della forma

$$P+\sqrt{D'}R=\left(it+\sqrt{D'}i\right)^n$$

per qualche $n\in\mathbb{N}$.

Scrivendole esplicitamente si ottiene che

$$R=i^n\sum_{\substack{k=0\\ k\text{ dispari}}}^{n}\sqrt{D'}^{k-1}t^{n-k}\binom{n}{k}$$

ma, poiché a è reale positivo, si ha $\frac{R(a)}{i^n}>0$, violando così l'esistenza di soluzioni.

Capitolo 4

Lo strumento delle Frazioni Continue

4.1 Frazioni Continue

Introdurremo qui lo strumento delle frazioni continue, seguendo la trattazione di van der Poorten e Tran [10] con alcune precisazioni tratte da [8]. Vedremo che le frazioni continue saranno molto utili per la risoluzione della nostra equazione.

Per quanto scritto in questa sezione considereremo da approssimare dei numeri $\alpha \in \mathbb{R}^+$ oppure $\alpha \in K((t^{-1}))$.

Un fatto curioso è che i teoremi di questa parte hanno lo stesso enunciato per i numeri interi e per i polinomi, mentre le dimostrazioni sono leggermente diverse (anche se di poco) nei due casi. Per questo le dimostrazioni verranno divise nei due casi in mezzo alla stessa, evitando di essere ripetute due volte.

4.1.1 Valori assoluti

Definizione 1: Un valore assoluto su un campo L è una funzione $|\cdot| : L \to \mathbb{R} \cup \{\infty\}$ con le seguenti proprietà:

- $\forall x \quad |x| \geq 0$
- $|x| = 0$ se e solo se $x = 0$
- $|xy| = |x||y|$
- $|x + y| \leq |x| + |y|$

Fatto 2: Per $L = \mathbb{R}$, l'usuale funzione valore assoluto è un valore assoluto.

Fatto 3: Per $L = K((t^{-1}))$ si può considerare la funzione grado come definita precedentemente e definire $|p| = e^{\deg p}$ con l'accortezza che, se $p = 0$, si ha $\deg p = -\infty$ e quindi $|p| = 0$.

L'unica proprietà non immediata da verificare è $|p + q| \leq |p| + |q|$ per la quale si può supporre senza perdita di generalità $\deg p > \deg q$ e quindi $\deg(p + q) \leq \deg p \implies |p + q| \leq |p| \leq |p| + |q|$.

Funzione di approssimazione

Abbiamo già definito, sia nel caso dei numeri interi sia nel caso delle serie di Laurent, una funzione di approssimazione $\lfloor\cdot\rfloor$. Essa corrisponde, nel caso degli interi, alla funzione "parte intera inferiore" e nel caso delle serie di Laurent, alla funzione "parte polinomiale".

Osserviamo ora che essa si comporta bene rispetto ai valori assoluti:

Osservazione 7. Dato $\alpha \in L$ si ha che $|\alpha - \lfloor\alpha\rfloor| < 1$ ed inoltre $|\lfloor\alpha\rfloor| \leq |\alpha|$.

La funzione $\lfloor\cdot\rfloor$ è inoltre l'unica con queste due proprietà (nei casi che consideriamo).

4.1.2 Definizione di Frazione Continua

Il nostro scopo sarà quello di prendere un elemento α e di cercare di approssimarlo attraverso altri elementi a_i.

Con la scrittura

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

o anche $\alpha = [a_0; a_1, a_2, \ldots]$ intendiamo la seguente cosa:

Riusciamo a definire una sequenza di razionali $\frac{x_n}{y_n} = [a_0; a_1, \ldots, a_n]$ ottenuti troncando la frazione continua infinita, con x_n ed y_n coprimi, in modo che, in un certo senso, $\frac{x_n}{y_n} \to \alpha$: se partiamo da un numero $\alpha \in \mathbb{R}^+$, gli a_i saranno numeri naturali e la convergenza sarà proprio quella di $\mathbb{R}$.

Nel caso invece dell'anello di serie di Laurent $\alpha \in K((t^{-1}))$ chiederemo elementi $a_i \in K[t]$ e la convergenza sarà quella delle migliori approssimazioni razionali, che vedremo a breve.

Definizione 2 (Resti e convergenti)**:** Per fare ciò ci serviremo della funzione di approssimazione $\lfloor\cdot\rfloor$: dato un elemento F definiamo per ricorsione la seguente successione:

$$F_0 = F, \qquad a_i = \lfloor F_i \rfloor, \qquad F_{i+1} = \frac{1}{F_i - a_i}$$

e, a partire da quest'ultima, definiamo anche le due successioni x_h, y_h:

$$x_{-1} = y_0 = 1, \qquad y_{-1} = 0, \qquad x_0 = a_0, \qquad y_{h+1} = a_h y_h + y_{h-1}, \qquad x_{h+1} = a_h x_h + x_{h-1}$$

Le F_i verranno chiamati resti della frazione continua di F, $\frac{x_h}{y_h}$ saranno invece dette convergenti.

Osservazione 8. Notiamo che definiti x_n ed y_n come sopra vale la relazione matriciale

$$\begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix}\cdots\begin{pmatrix} a_h & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} x_h & x_{h-1} \\ y_h & y_{h-1} \end{pmatrix} \tag{4.1}$$

Dimostrazione. Tutto ciò può essere controllato per induzione su h ricordando la definizione

$$[a_0; a_1, \ldots, a_h] = a_0 + \frac{1}{[a_1; a_2, \ldots, a_h]}, \qquad [a_0] = a_0$$

□

Siccome le F_i definite sopra soddisfano la relazione $F = [a_0, a_1, \ldots, a_h, F_{h+1}]$ abbiamo dalla corrispondenza che

$$\begin{aligned} F \leftrightarrow & \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix}\cdots\begin{pmatrix} a_h & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} F_{h+1} & 1 \\ 1 & 0 \end{pmatrix} \\ = & \begin{pmatrix} x_h & x_{h-1} \\ y_h & y_{h-1} \end{pmatrix}\begin{pmatrix} F_{h+1} & 1 \\ 1 & 0 \end{pmatrix} \leftrightarrow \frac{x_h F_{h+1} + x_{h-1}}{y_h F_{h+1} + y_{h-1}} \end{aligned}$$

ovvero otteniamo

$$F = \frac{x_h F_{h+1} + x_{h-1}}{y_h F_{h+1} + y_{h-1}}, \qquad F_{h+1} = -\frac{y_{h-1}F - x_{h-1}}{y_h F - x_h}$$

da cui, siccome il prodotto è telescopico, risulta che

$$(-1)^{h+1} F_1 F_2 \ldots F_{h+1} = (y_h F - x_h)^{-1}$$

Notiamo inoltre che trasponendo la relazione matriciale si ottiene

$$\frac{y_h}{y_{h-1}} = [a_h, a_{h-1}, \ldots, a_1]$$

e prendendone i determinanti si ha

$$x_h y_{h-1} - x_{h-1} y_h = (-1)^{h-1}$$

Per induzione è inoltre abbastanza semplice dimostrare

$$x_k y_{k-2} - y_k x_{k-2} = (-1)^{k-1} a_k$$

Inoltre sostitudendo F come ricavato sopra da F_{h+1} si ha

$$y_h F - x_h = \frac{y_h x_{h-1} - y_{h-1} x_h}{y_h F_{h+1} + y_{h-1}} = \frac{(-1)^h}{y_h} \cdot \frac{1}{(F_{h+1} + \frac{y_{h-1}}{y_h})}$$

che mostra la qualità eccellente delle approssimazioni di F data dalle sue convergenti. In particolare, passando ai valori assoluti, si ha che

$$|y_h F - x_h| = \frac{1}{|y_h|\left|F_{h+1} + \frac{y_{h-1}}{y_h}\right|} < \frac{1}{|y_h|}$$

dove l'ultima minorazione segue osservando che nel caso dei polinomi $\deg F_{h+1} > \deg \frac{y_{h-1}}{y_h}$ e nel caso degli interi si ha $F_{h+1} > 1$ e $\frac{y_{h-1}}{y_h} > 0$ (visto che $\alpha \in \mathbb{R}^+$).

Lemma 7: Si ha che $|y_h| > |y_{h-1}|$ ed inoltre $|y_n| \to \infty$ in n.

Dimostrazione. Per la prima parte osserviamo che la tesi è vera per $h = 1$ e si ha inoltre $|a_i| \geq 1$ da cui segue che

$$|y_{h+1}| = |a_h y_h + y_{h-1}| > |a_h||y_h| \geq |y_h|$$

nel caso dei reali. Nel caso dei polinomi si ha invece per ipotesi induttiva che $\deg(a_h y_h) > \deg(y_{h-1})$ e quindi si ottiene nuovamente la tesi.

La seconda parte segue ricordando che, in entrambi i casi, $|y_n|$ può assumere un numero discreto di valori il cui unico punto di accumulazione è all'infinito. (Nel caso dei reali ciò segue poiché $y_n \in \mathbb{N}$). □

Lemma 8: Per i numeri reali, le convergenti pari sono sempre maggiori delle convergenti dispari. Inoltre le convergenti pari formano una sequenza decrescente e le convergenti dispari formano una sequenza crescente che tende ad α, numero della frazione continua.

Dimostrazione. Detta $C_k = \frac{x_k}{y_k}$ sappiamo, utilizzando le formule appena dimostrate, che

$$\forall k \geq 3 \quad C_k - C_{k-2} = \frac{(-1)^{k-1} a_k}{y_k y_{k-2}}$$

$$\forall k \geq 2 \quad C_k - C_{k-1} = \frac{(-1)^k}{y_k y_{k-1}}$$

dove gli y_k sono i denominatori delle approssimanti. Siccome gli $a_k > 0$ si ha che le convergenti pari sono decrescenti e le convergenti dispari sono decrescenti.

Inoltre dalla seconda equazione si ottiene che quelle pari sono sempre maggiori delle dispari come desiderato. □

Il lemma che segue fornisce un inverso a quanto osservato poc'anzi, mostrando che se si ha una approssimazione particolarmente buona dell'elemento F, essa deve essere una convergente in frazione continue.

Lemma 9: Siano $x, y \in R$ relativamente coprimi. Allora si ha

$$|yF - x| < \frac{1}{|y|}$$

se e solo se la frazione $\frac{x}{y}$ è una convergente di F.

Dimostrazione (Per gli interi)**.** Il fatto che la disuguaglianza valga se x/y è una convergente è stato dimostrato poco sopra.

Supponiamo allora che vi sia una frazione x/y per la quale si ha la disuguaglianza di cui sopra. Consideriamo h tale che $y_h \leq y < y_{h+1}$.

Supponiamo inoltre che non valga $x = x_h$ e $y = y_h$ nel qual caso è banalmente vero. Risolvendo per a e b il sistema di equazioni seguente

$$\begin{aligned} y &= a y_{h-1} + b y_h \\ x &= a x_{h-1} + b x_h \end{aligned}$$

otteniamo che $a = (-1)^h (x_h y - y_h x)$ e $b = (-1)^h (x y_{h-1} - x_{h-1} y)$ e quindi a e b sono interi, entrambi non nulli (altrimenti $\frac{x}{y}$ sarebbe una convergente).

Ora, siccome $0 < y = a y_{h-1} + b y_h < y_h$ si ha che a e b devono avere segni opposti. Si ha quindi che $a(y_{h-1} F - x_{h-1})$ e $b(y_h F - x_h)$ hanno lo stesso segno (perché le convergenti

sono alternatamente più grandi e più piccole del valore limite), da cui segue:

$$\begin{aligned} |yF - x| &= |(ay_{h-1} + by_h)F - (ax_{h-1} + bx_h)| \\ &= |a(y_{h-1}F - x_{h-1}) + b(y_hF - x_h)| \\ &= |a||y_{h-1}F - x_{h-1}| + |b||y_hF - x_h| \\ &> |b||y_hF - x_h| \\ &\geq |y_hF - x_h| \end{aligned}$$

confermando il fatto che le convergenti sono le approssimazioni localmente migliori. □

Dimostrazione (Per i polinomi)**.** La parte del se è stata osservata poco sopra. Possiamo allora prendere h tale che $\deg y_{h-1} \leq \deg y < \deg y_h$ e notare che supporre che $\frac{x}{y}$ non sia una convergente ci dice che y non è un multiplo costante di y_{h-1}. Visto che $x_hy_{h-1} - x_{h-1}y_h = \pm 1$, esistono degli elementi non nulli $a \in R$ e $b \in R$ (basta risolvere il sistema lineare sottostante invertendo la matrice) tali che:

$$\begin{aligned} y &= ay_{h-1} + by_h \\ x &= ax_{h-1} + bx_h \end{aligned}$$

e quindi $yF - x = a(y_{h-1}F - x_{h-1}) + b(y_hF - x_h)$. Supponiamo adesso che i due termini sulla destra abbiano valore assoluto diverso, $\deg a - \deg y_h$ e $\deg b - \deg y_{h+1}$, rispettivamente. In questo caso ovviamente $\deg(yF-x) > \deg(y_{h-1}F-x_{h-1}) > \deg(y_hF-x_h)$, confermando che le convergenti danno l'approssimazione localmente migliore di F. Ed otteniamo quindi l'assurdo utilizzando che $\deg(y_{h-1}F - x_{h-1}) = -\deg(y_h)$.

Per verificare che siano effettivamente differenti i due gradi, osserviamo che $\deg ay_{h-1} = \deg by_h$, altrimenti non può essere che $\deg y < \deg y_h$.

Allora si ha $\deg a - \deg y_h = \deg b - \deg y_{h-1} > \deg b - \deg y_{h+1}$. Inoltre $\deg a - \deg y_h = \deg(yF - x)$. Di più, siccome $\deg a \geq \deg y_h - \deg y_{h-1}$, è ovvio che $\deg a - \deg y_h \geq -\deg y$. □

4.1.3 Legami con la Pell e radice quadrata

Come abbiamo visto, le convergenti alla frazione continua sono buone approssimazioni del numero da cui si parte. Guidati da ciò, se proviamo a dividere per y^2 la 1.1 con $c = 1$ otteniamo che

$$\left(\frac{x}{y}\right)^2 - d = \frac{1}{y^2}$$

il che ci porta a pensare che $\frac{x}{y}$ debba essere una buona approssimazione di $\sqrt{d}$.

In particolare si ha il seguente lemma:

Lemma 10: Sia $\frac{x}{y} > 0$ tale che $\left(\frac{x}{y}\right)^2 - D = \frac{1}{y^2}$. Allora $\frac{x}{y}$ è convergente a $\sqrt{D}$

Dimostrazione. Basta dimostrare che $\left|\frac{x}{y} - \sqrt{D}\right| < \frac{1}{|y^2|}$. Ciò segue in quanto per ipotesi

$$\left(\frac{x}{y} - \sqrt{D}\right)\left(\frac{x}{y} + \sqrt{D}\right) = \frac{1}{y^2}$$

ma si ha anche che $\frac{x}{y} + \sqrt{D} \geq \sqrt{D}$ e quindi

$$\left(\frac{x}{y} - \sqrt{D}\right) \leq \frac{1}{\sqrt{D}y^2} < \frac{1}{y^2}$$

□

Osservazione 9. Da quanto appena dimostrato segue che se $c < \sqrt{D}$ ogni soluzione della Pell $x^2 - Dy^2 = c$ è data da una convergente a $\sqrt{D}$.

Definiamo quindi $\sqrt{d}$ anche per i polinomi per verificare poi la correttezza della nostra intuizione.

Radice quadrata di una serie di Laurent

Data una serie di laurent $D \in K((t^{-1}))$ diciamo che $F \in K((t^{-1}))$ è una sua radice quadrata se si ha $F \cdot F = D$. Notiamo che, affinché essa esista, è necessario e sufficiente che il coefficiente di testa di D sia un quadrato in K e che il grado di D sia pari, le stesse condizioni che erano necessarie affinché esistesse una soluzione non banale all'equazione di Pell per D.

Basta infatti impostare la generica $\left(\sum_{h=-k}^{\infty} f_h t^{-h}\right)^2 = \sum_{h=-r}^{\infty} d_h t^{-h}$ per rendersi conto delle condizioni ed in particolare per ottenere delle formule esplicite per il calcolo della radice quadrata, ovvero $\sum_{i+j=h} f_i f_j = d_h$, da cui si ricava

$$f_{-\frac{r}{2}} = \sqrt{d_{-r}}$$

$$f_{h-\frac{r}{2}} = \frac{1}{2f_{-\frac{r}{2}}}\left(d_{h-r} - \sum_{\substack{i+j=h-r \\ i,j>-\frac{r}{2}}} f_i f_j\right) \quad \text{per } h > 0$$

4.1.4 Frazione continua della radice quadrata

Lemma 11: Mostriamo ora per induzione che tutti i resti della frazione continua della radice quadrata di un elemento $D \in R$ sono della forma

$$F_i = \frac{P_i + \sqrt{D}}{Q_i}$$

dove $P_i, Q_i \in R$ e sono tali che $Q_i \mid D - P_i^2$

Dimostrazione. Banalmente si ha che $F = F_0$ con $P_0 = 0$ e $Q_0 = 1$.

Osserviamo inoltre che $a_i = \lfloor F_i \rfloor = \left\lfloor \frac{P_i + \sqrt{D}}{Q_i} \right\rfloor = \left\lfloor \frac{P_i + \lfloor \sqrt{D} \rfloor}{Q_i} \right\rfloor$ e coincide con il quoziente della divisione euclidea di $P_i + \left\lfloor \sqrt{D} \right\rfloor$ per Q_i.

Allora semplicemente effettuando il passaggio dell'algoritmo per lo sviluppo in frazione continua otteniamo le formule per P_{i+1} e Q_{i+1} noti solo $P_i, Q_i, E = \left\lfloor \sqrt{D} \right\rfloor$ ed a_i.

In definitiva abbiamo le seguenti formule per il calcolo della frazione continua della radice quadrata:

$$F_r = \frac{P_r + \sqrt{D}}{Q_r} \qquad F_{r+1} = \frac{1}{F_r - a_r} \tag{4.2}$$

$$a_r = \left\lfloor \frac{P_r + E}{Q_r} \right\rfloor \qquad E = \left\lfloor \sqrt{D} \right\rfloor \tag{4.3}$$

$$P_{r+1} = a_r Q_r - P_r \qquad Q_{r+1} = \frac{D - P_{r+1}^2}{Q_r} \tag{4.4}$$

$$P_0 = 0 \qquad Q_0 = 1 \tag{4.5}$$

$$\begin{cases} x_{h+1} = a_h x_h + x_{h-1} \\ y_{h+1} = a_h y_h + y_{h-1} \end{cases}$$

dove $x_{-1} = 1$, $y_{-1} = 0$, $x_0 = 0$, $y_0 = 1$. □

Osservazione 10. Detto L il campo delle frazioni di R si ha che per $D \in R$, tutti i resti di $\sqrt{D}$ sono contenuti in $L(\sqrt{D})$.

Esempio 2: Calcoliamo i primi termini della frazione continua della radice quadrata di $D = t^6 + t$. Con un veloce conto si scopre che $E = \left\lfloor \sqrt{D} \right\rfloor = t^3$.

i	P_i	Q_i	a_i	x_i	y_i
-1	/	/	/	1	0
0	0	1	t^3	t^3	1
1	t^3	t	$2t^2$	$2t^5+1$	$2t^2$
2	t^3	$\boxed{1}$	$2t^3$	/	/
3	t^3	t	$2t^2$	/	/

Come vedremo più avanti avere $Q_r = 1$ ci indica che la frazione continua è periodica (questo lo potevamo vedere anche osservando che $P_3 = P_1$ e $Q_3 = Q_1$ e, siccome tutto l'algoritmo dipende solo dai Q_i e P_i, si sarebbe dovuto ripetere) e quindi abbiamo ottenuto

$$\sqrt{t^6 + t} = [t^3; 2t^2, 2t^3, 2t^2, 2t^3, \ldots]$$

In particolare, visto che il nostro claim è che le soluzioni della Pell debbano essere buone approssimazioni di $\sqrt{D}$, possiamo provare con le convergenti, calcolando $x_i^2 - Dy_i^2$, ottenendo:

$$x_0^2 - Dy_0^2 = -t, \qquad x_1^2 - Dy_1^2 = 1$$

e quindi abbiamo effettivamente risolto l'equazione di Pell con $D = t^6 + t$.

4.2 Elementi Ridotti e Frazioni Continue

4.2.1 Frazioni Continue Periodiche

Definizione 3 (Periodicità)**:** Data una frazione continua $\alpha = [a_0; a_1, a_2, \ldots]$ diciamo che essa è periodica se $\exists k \in \mathbb{N}$ tale che $\forall r \geq r_0$ si ha $a_{r+k} = a_r$, ovvero se la stringa degli a_i è definitivamente periodica.

Definizione 4 (Pura Periodicità)**:** Data una frazione continua $\alpha = [a_0; a_1, a_2, \ldots]$ diciamo che essa è puramente periodica se la definizione precedente vale con $r_0 = 0$.

Per avere una notazione più pulita nel seguito definiamo $\iota(n) = (-1)^n$.

Definizione 5 (Quasi-Periodicità)**:** Una frazione continua $\alpha = [a_0; a_1, a_2, \ldots]$ si dice quasi-periodica se $\exists l \in \mathbb{N}$, $\exists m \in \mathbb{N}$, $\exists \mu \in R^*$ tale che

$$\forall n \geq m \quad a_n = \mu^{\iota(n)} a_{n+l}$$

ovvero se è definitivamente periodica a meno di moltiplicazione per un invertibile fissato.

Se $m = 0$ si dice puramente quasi-periodica.

Criterio 12 (Per la periodicità)**:** Controllare se una frazione continua è periodica consta di un numero infinito di condizioni. Non è però difficile verificare che, siccome la coda della frazione continua è completamente determinata dal resto n-esimo, si ha:

- CF(α) è periodica se e solo se $\exists k \in \mathbb{N}$ tale che $\exists m \in \mathbb{N} \quad F_{k+m} = F_m$
- CF(α) è puramente periodica se e solo se $\exists k \in \mathbb{N}$ tale che $\alpha = F_k$ (visto che $F_0 = \alpha$ per definizione)

dove F_k indica il resto k-esimo, ovvero $\alpha = [a_0, \ldots, a_{k-1}, F_k]$.

4.2.2 CF($\sqrt{D}$) e la periodicità

Vogliamo capire meglio l'espansione in frazioni continue di $\sqrt{D}$. Vedremo che si può solitamente andare all'indietro nell'algoritmo delle frazioni continue, ovvero si possono avere solo pre-periodi corti.

Nel nostro caso, come visto precedentemente, tutti i resti sono contenuti nell'estensione quadratica $L[\sqrt{D}]$ di L (che nel nostro caso è rispettivamente $\mathbb{Q}$ oppure $K(t)$ e ricordiamo che R è rispettivamente $\mathbb{Z}$ o $K[t]$). Essa ha esattamente un automorfismo σ su L non banale, ovvero quello tale che $\sigma(\sqrt{D}) = -\sqrt{D}$.

Definizione 6: $\alpha \in K(t, \sqrt{D})$ si dice σ-ridotto se

$$\deg \sigma(\alpha) > 0 > \deg \alpha$$

dove il grado è inteso nell'immersione canonica $K(t, \sqrt{D}) \subseteq K((t^{-1}))$

Definizione 7: $\alpha \in \mathbb{Q}(\sqrt{D})$ si dice σ-ridotto se

$$\alpha > 1, \qquad -1 < \sigma(\alpha) < 0$$

dove l'ordine è inteso rispetto all'immersione canonica $\mathbb{Q}(\sqrt{D}) \subseteq \mathbb{R}$

Lemma 13: Sia $\alpha \in L(\sqrt{D}) \setminus L$. Allora esiste al più un $a \in S$ tale che $|a| \leq |\alpha|$ e $a + \alpha$ è σ-ridotto.

Dimostrazione. Si ha $\sigma(a+\alpha) = a+\sigma(\alpha)$ da cui se vogliamo $|\sigma(a+\alpha)| = |a+\alpha| > 1$ deve essere $a = -\lfloor\alpha\rfloor$ per il lemma sull'unicità. □

Proposizione 14: Se il resto F_m è σ-ridotto, lo è anche F_n per ogni $n \geq m$.

Dimostrazione (Per i Polinomi)**.** Basta mostrarlo per $n = m+1$: Sappiamo ovviamente che per ogni F_n si ha $\deg F_n > 0$ ed inoltre

$$\sigma(F_{m+1}) = \frac{1}{\sigma(F_m) - a_m}$$

e $|a_m| \leq |F_m| < |\sigma(F_m)|$ implica che $\deg \sigma(F_{m+1}) = -\deg F_m > 0$ □

Dimostrazione (Per gli Interi)**.** Basta anche in questo caso mostrarlo per $n = m+1$: Sappiamo che per ogni F_n si ha $F_n > 1$ ed anche

$$\sigma(F_{m+1}) = \frac{1}{\sigma(F_m) - a_m}$$

con $\sigma(F_m) < 0 < a_m$ e quindi $\sigma(F_{m+1}) < 0$. Inoltre affinché si abbia $\sigma(F_{m+1}) > -1$ deve essere $1 < a_m - \sigma(F_m)$ ma ciò è banalmente vero siccome $\sigma(F_m) < 0$ e $a_m \geq 1$ poiché intero. □

Lemma 15: α è σ-ridotto se e solo se $-\frac{1}{\sigma(\alpha)}$ è σ-ridotto.

Dimostrazione (Per i Polinomi)**.** Conseguenza immediata delle due formule

$$\deg \alpha = -\deg\left(\sigma\left(\frac{-1}{\sigma(\alpha)}\right)\right), \qquad \deg \sigma(\alpha) = -\deg\left(\frac{-1}{\sigma(\alpha)}\right)$$

□

Dimostrazione (Per gli interi)**.** Notiamo che l'operazione $\alpha \mapsto -\frac{1}{\sigma(\alpha)}$ è involutiva e basta quindi mostrare una sola freccia. Tutte le disuguaglianze seguono attraverso semplici manipolazioni algebriche e non vengono quindi eseguite. □

Proposizione 16: Supponiamo che $\alpha_1 \in L(\sqrt{D})$ sia σ-ridotto. Allora esiste un unico $\alpha_0 \in L(\sqrt{D})$ che è σ-ridotto e soddisfa

$$\alpha_1 = \frac{1}{\alpha_0 - \lfloor\alpha_0\rfloor}$$

Dimostrazione. Sappiamo che esiste al più un $a_0 \in R$ tale che $\alpha_0 = a_0 + \frac{1}{\alpha_1}$ è σ-ridotto, ovvero $a_0 = \left\lfloor\frac{-1}{\sigma(\alpha_1)}\right\rfloor$ riscrivendolo come

$$\frac{-1}{\sigma(\alpha_0)} = \frac{1}{\frac{-1}{\sigma(\alpha_1)} - a_0}$$

vediamo che α_0 è σ-ridotto applicando il lemma di cui sopra. Inoltre, visto che $\deg \alpha_1 < 0$ è chiaro che $a_0 = \lfloor\alpha_0\rfloor$. □

Lemma 17: Supponiamo che α_m sia σ-ridotto e $\mathrm{CF}(\alpha_m)$ sia (quasi-)periodica, allora $\mathrm{CF}(\alpha_m)$ è puramente (quasi-)periodica

Dimostrazione. Supponiamo $n > m$, $l \in \mathbb{N}$ e $\mu \in R^*$ (con $\mu = 1$ nel caso di periodicità) con $\alpha_n = \mu^{\iota(n)}\alpha_{n+l}$. Per la proposizione precedente, $\alpha_{n-1}, \alpha_n, \alpha_{n+l-1}, \alpha_{n+l}$ sono tutte σ-ridotte e quindi abbiamo

$$\alpha_n = \frac{1}{\alpha_{n-1} - a_{n-1}} = \mu^{\iota(n)}\alpha_{n+l} = \mu^{\iota(n)}\frac{1}{\alpha_{n+l-1} - a_{n+l-1}} = \frac{1}{\mu^{\iota(n-1)}\alpha_{n+l-1} - \mu^{\iota(n-1)}a_{n+l-1}}$$

con $\left\lfloor \mu^{\iota(n-1)}\alpha_{n+l-1} \right\rfloor = \mu^{\iota(n-1)}a_{n+l-1}$, la proposizione precedente implica che $\alpha_{n-1} = \mu^{\iota(n-1)}\alpha_{n+l-1}$ come desiderato, e possiamo ripetere questo ragionamento fino ad arrivare a $\alpha_m = \mu^{\iota(m)}\alpha_{m+l}$. □

Osservazione 11 ($\sqrt{D}$ ha pre-periodo di lunghezza 1)**.** $\left\lfloor \sqrt{D} \right\rfloor + \sqrt{D}$ è σ-ridotto e la sua espansione in frazione continua differisce da quella di $\sqrt{D}$ solo per il primo termine. Quindi se $\sqrt{D}$ ha espansione periodica, allora $\sqrt{D}$ ha pre-periodo uno.

Osservazione 12 (Sviluppo di $\sqrt{D}$)**.** Se $\sqrt{D}$ ha espansione periodica, vale

$$\sqrt{D} = [a_0, a_1, \ldots, a_n, \overline{2a_0, a_1, \ldots, a_n}]$$

4.3 Irrazionali quadratici e frazioni continue periodiche

4.3.1 Frazioni continue periodiche

Lemma 18: La frazione continua di $\sqrt{D}$ è periodica se e solo se esiste r tale che $Q_r = 1$.

Dimostrazione. $\boxed{\Leftarrow}$ Se $Q_r = 1$, si può vedere eseguendo i calcoli come visto precedentemente che $P_{r+1} = P_1$ e $Q_{r+1} = Q_1$, il che ci dà quindi la periodicità
$\boxed{\Rightarrow}$ Abbiamo visto nella sezione precedente che $\sqrt{D}$ è periodica se e solo se $\left\lfloor \sqrt{D} \right\rfloor + \sqrt{D}$ è puramente periodica, il che ci dice che $\sqrt{D}$ può avere pre-periodo di lunghezza uno al più. Quindi se è periodica vuol dire che esiste k tale che si ha $F_1 = F_{k+1}$, ovvero si ha $P_{k+1} = P_1 = \left\lfloor \sqrt{D} \right\rfloor$ e $Q_{k+1} = Q_1 = D - \left\lfloor \sqrt{D} \right\rfloor^2$. Allora si ottiene che

$$Q_k = \frac{D - P_{k+1}^2}{Q_{k+1}} = 1$$

che ci dà la tesi desiderata. □

4.3.2 Convergenti alla radice quadrata ed equazione di Pell

Motivati dal legame tra frazioni continue ed equazione di Pell, ci si potrebbe chiedere effettivamente "quanto bene" le convergenti a $\sqrt{D}$ risolvano l'equazione di Pell. Significativamente si ottiene:

Lemma 19: Consideriamo lo sviluppo in frazione continua di $\sqrt{D}$: allora, per ogni h vale

$$x_h^2 - Dy_h^2 = (-1)^{h+1}Q_{h+1}$$

Dimostrazione. Basta ricordarsi la formula telescopica

$$(-1)^{h+1}F_1F_2 \ldots F_{h+1} = (y_hF - x_h)^{-1}$$

e considerarne la norma:

$$\mathcal{N}(F_s) = \frac{P_s^2 - D}{Q_s^2} = -\frac{Q_s Q_{s-1}}{Q_s^2} = -\frac{Q_{s-1}}{Q_s}$$

e si ottiene quindi

$$(-1)^{h+1}\frac{Q_0}{Q_1} \cdot \frac{Q_1}{Q_2} \cdot \ldots \cdot \frac{Q_h}{Q_{h+1}}\mathcal{N}(y_h F - x_h) = 1$$

e ricordando che $F = \sqrt{D}$ si ottiene quindi la tesi. □

Abbiamo già visto che se l'equazione di Pell ammette soluzioni non banali, allora esse sono convergenti alla frazione continua di $\sqrt{D}$. Dal lemma precedente, se una convergente risolve l'equazione di Pell, deve essere $Q_{h+1} = \pm 1$. Se si ottiene $Q_{h+1} = -1$, come già notato precedentemente, basta elevare la soluzione al quadrato per ottenerne una con $Q_{h+1} = 1$.

Per la caratterizzazione precedente della periodicità delle frazioni continue otteniamo quindi un noto teorema di Abel:

Teorema 20 (di Abel di caratterizzazione dei polinomi Pelliani)**:** L'equazione di Pell per il polinomio $D(t) \in K[t]$

$$P(t)^2 - D(t)Q(t)^2 = 1$$

è risolubile in $K[t]$ se e solo se lo sviluppo in frazione continua di $\sqrt{D(t)}$ è periodico.

Dimostrazione. Basta ora unire i lemmi precedenti:

- Se l'equazione è risolubile sappiamo che la soluzione $\frac{P(t)}{Q(t)}$ deve essere un'approssimante di $\sqrt{D}$
- Dal lemma precedente sappiamo che se è l'approssimante h-esima, allora si ha $Q_{h+1} = \pm 1$. (Nel caso sia -1 si può semplicemente elevare al quadrato la soluzione).
- Visto che $\exists r : Q_r = 1$ la frazione continua di $\sqrt{D(t)}$ è periodica.

Viceversa, se la frazione continua è periodica, allora per il lemma precedente esiste $Q_r = 1$ e quindi si ha

$$x_r^2 - D y_r^2 = \pm 1$$

e basta elevare al quadrato la soluzione per ottenere una soluzione alla Pell con $c = 1$. □

Come vedremo a breve vi è una sostanziale differenza nel caso dei polinomi o nel caso degli interi: in quest'ultimo infatti tutte le equazioni di Pell con $c = 1$ sono risolubili, sostanzialmente perché mostreremo che ogni frazione continua di un numero irrazionale quadratico è periodica. Nel caso dei polinomi su campi infiniti non si ha invece questo risultato, ed anzi le soluzioni sono "rare".

4.3.3 Ogni irrazionale quadratico ha frazione continua periodica

La dimostrazione che segue è presa da [5] di un bellissimo teorema di Lagrange che lega gli irrazionali quadratici alle frazioni continue periodiche.

In analogia con quanto succede per le frazioni decimali, possiamo indicare una tale frazione continua periodica come segue

$$\alpha = [a_0; a_1, a_2, \ldots, a_{k_0-1}, \overline{a_{k_0}, a_{k_0+1}, \ldots, a_{k_0+h-1}}]$$

Teorema 21: Sia $\alpha \in \mathbb{R}^+$. Allora la frazione continua di α è periodica se e solo se α è un irrazionale quadratico.

Dimostrazione. $\boxed{\Rightarrow}$ Ovviamente i resti delle frazioni continue periodiche soddisfano la relazione

$$r_{k+h} = r_k \qquad (k \geq k_0)$$

Allora, sulla base della formula per i resti si ha, per $k \geq k_0$:

$$\alpha = \frac{p_{k-1}r_k + p_{k-2}}{q_{k-1}r_k + q_{k-2}} = \frac{p_{k+h-1}r_k + p_{k+h-2}}{q_{k+h-1}r_k + q_{k+h-2}}$$

di modo che si abbia

$$\frac{p_{k-1}r_k + p_{k-2}}{q_{k-1}r_k + q_{k-2}} = \frac{p_{k+h-1}r_k + p_{k+h-2}}{q_{k+h-1}r_k + q_{k+h-2}}$$

e quindi il numero r_k soddisfa una equazione quadratica a coefficienti interi e quindi è un irrazionale quadratico. Ma per la quindi per la prima uguaglianza si ha che anche α è un irrazionale quadratico.

$\boxed{\Leftarrow}$ Il viceversa è un po' più complicato: supponiamo che il numero α soddisfi un'equazione quadratica a coefficienti interi

$$a\alpha^2 + b\alpha + c = 0$$

Se scriviamo α in termini dei suoi resti di ordine n abbiamo

$$\alpha = \frac{p_{n-1}r_n + p_{n-2}}{q_{n-1}r_n + q_{n-2}}$$

e quindi otteniamo che gli r_n soddisfano l'equazione

$$A_n r_n^2 + B_n r_n + C_n = 0$$

dove A_n, B_n, C_n sono interi definiti da

$$\begin{aligned} A_n &= ap_{n-1}^2 + bp_{n-1}q_{n-1} + cq_{n-1}^2 \\ B_n &= 2ap_{n-1}p_{n-2} + b(p_{n-1}q_{n-2} + p_{n-2}q_{n-1}) + 2cq_{n-1}q_{n-2} \\ C_n &= ap_{n-2}^2 + bp_{n-2}q_{n-2} + cq_{n-2}^2 \end{aligned}$$

dalle quali segue in particolare che $C_n = A_{n-1}$.

Con queste formule si può facilmente verificare che

$$B_n^2 - 4A_nC_n = (b^2 - 4ac)(p_{n-1}q_{n-2} - q_{n-1}p_{n-2})^2 = b^2 - 4ac$$

e quindi il discriminante dell'equazione è lo stesso per tutti gli n ed è uguale al discriminante dell'equazione che avevamo per α.

Inoltre, siccome si ha

$$\left|\alpha - \frac{p_{n-1}}{q_{n-1}}\right| < \frac{1}{q_{n-1}^2}$$

ne segue che

$$p_{n-1} = \alpha q_{n-1} + \frac{\delta_{n-1}}{q_{n-1}} \qquad (|\delta_{n-1}| < 1)$$

e quindi, la prima formula esplicita per gli A_n ci dice che

$$\begin{aligned} A_n &= a\left(aq_{n-1} + \frac{\delta_{n-1}}{q_{n-1}}\right)^2 + b\left(\alpha q_{n-1} + \frac{\delta_{n-1}}{q_{n-1}}\right)q_{n-1} + cq_{n-1}^2 \\ &= \big(a\alpha^2 + b\alpha + c\big)q_{n-1}^2 + 2a\alpha\delta_{n-1} + a\frac{\delta_{n-1}^2}{q_{n-1}^2} + b\delta_{n-1} \end{aligned}$$

e ci dà

$$|A_n| = \left|2a\alpha\delta_{n-1} + a\frac{\delta_{n-1}^2}{q_{n-1}} + b\delta_{n-1}\right| < 2|a\alpha| + |a| + |b|$$

e abbiamo anche

$$|C_n| = |A_{n-1}| < 2|a\alpha| + |a| + |b|$$

Quindi, i coefficienti A_n e C_n sono limitati in valore assoluto e quindi possono assumere solo un numero finito di valori distinti. Segue quindi sulla base dell'equazione del discriminante conservato che anche B_n può prendere solo un numero finito di valori distinti. Quindi, per n che cresce da 1 ad ∞, incontriamo solo un numero finito di equazioni distinte per r_n, che può prendere solo un numero finito di valori per ciascuna di queste equazioni.

Ne segue che, per k ed h appropriatamente scelti si ha

$$r_k = r_{k+h}$$

e ciò mostra che la frazione continua che rappresenta α è periodica e quindi mostra la tesi. □

Corollario 22: $d \in \mathbb{Z}$ non è un quadrato perfetto se e solo esistono soluzioni non banali all'equazione $x^2 - dy^2 = 1$

Dimostrazione. Se d è un quadrato perfetto, abbiamo già mostrato nell'osservazione 3 che le uniche soluzioni sono banali.

Se d non è un quadrato perfetto, per il teorema di Lagrange l'espansione in frazione continue di $\sqrt{d}$ è periodica e quindi per il teorema di Abel, l'equazione ammette una soluzione non banale. □

4.4 Risoluzione dell'Equazione con le Frazioni Continue

Ricapitoliamo brevemente come si risolve l'equazione di Pell con $c = 1$ utilizzando il metodo delle frazioni continue

- Si computa lo sviluppo in frazioni continue di $\sqrt{d}$. Se non periodico (cosa che può succedere solo per i polinomi su campi infiniti) per il Teorema di Abel l'equazione non ha soluzione.

- Se periodico si deve avere che $\exists r$ t.c. $Q_r = 1$ e quindi si ha $x_h^2 - dy_h^2 = (-1)^{h+1}Q_h = (-1)^{h+1}$ da cui si ottiene la soluzione fondamentale (eventualmente elevando al quadrato).

Capitolo 5

Argomenti Collegati

5.1 Unità di norma negativa

Come abbiamo visto la Pell ci dice che $\mathcal{O}_K$ ha delle unità non banali di norma positiva. Analogamente è molto importante riuscire a sapere anche se $\mathcal{O}_K$ ha delle unità di norma negativa.

Da Mollin e Srinivasan [9] abbiamo il seguente criterio: supponendo di avere la soluzione al caso della Pell con $c = 1$ riusciamo a sapre se anche $c = -1$ è risolubile oppure no. Ciò può sembrare la soluzione del problema, ma in realtà una risposta soddisfacente sarebbe ottenere una classificazione sensata dei d per cui la Pell con $c = -1$ ha soluzione. Il modo che esponiamo di seguito prevede di conoscere la soluzione alla Pell con $c = 1$, il che richiede un tempo abbastanza lungo e pertanto non viene considerato essere soddisfacente.

Teorema 23 (Unità di norma negativa)**:** Sia $(x_0, y_0) \in \mathbb{Z}^2$ la soluzione fondamentale alla Pell per $c = 1$ a d fissato. Allora $\exists x, y$ t.c. $x^2 - dy^2 = -1$ se e solo se $x_0 \equiv -1 \pmod{2d}$

Dimostrazione. $\boxed{\Rightarrow}$ Sia $(x, y) \in \mathbb{Z}^2$ la soluzione minima di norma negativa. Elevando al quadrato l'equazione otteniamo che $(x + \sqrt{d}y)^2 = x^2 + dy^2 + \sqrt{d}2xy$ è una soluzione alla Pell con $c = 1$. Come spiegheremo poco sotto, essa è necessariamente la minima. Si ha allora che $x_0 = x^2 + dy^2 = 2dy^2 - 1 \equiv -1\,(2d)$.

Vediamo perché è la minima: usando il teorema sulla forma delle soluzioni otteniamo che $\mathrm{Sol}_{\{\pm 1\}}$ è generato da una sola unità fondamentale (che è $x + \sqrt{d}y$) a meno di cambi di segno. L'unità fondamentale generante Sol_1 deve essere un suo multiplo intero a meno del segno. Ma già elevata al quadrato essa ci fornisce un elemento di Sol_1, che quindi deve essere la soluzione minima (per positività).

$\boxed{\Leftarrow}$ Supponiamo che si abbia $x_0 = 2da - 1$. Per la pell vale $dy_0^2 = x_0^2 - 1 = (x_0 + 1)(x_0 - 1) = 4da(da - 1)$ che implica che $y_0 = 2b$. Semplificando si ottiene $b^2 = a(da - 1)$. Siccome si ha $\mathrm{MCD}(a, da - 1) = \mathrm{MCD}(a, -1) = 1$ abbiamo che $a = c^2$ e $e^2 = da - 1 = dc^2 - 1$, ovvero e e c risolvono l'equazione di Pell per $c = -1$. □

5.2 Equazione di Pell generica

In questa sezione trattiamo la struttura ed il computo delle soluzioni dell'equazione di Pell 1.1 con c generico.

5.2.1 Struttura delle soluzioni

Mostriamo la struttura delle soluzioni dell'equazione $x^2 - Dy^2 = N$. Le idee di questa parte sono prese da [7].

Definizione 8 (Soluzione primitiva)**:** Una soluzione di $x^2 - Dy^2 = N$ viene detta primitiva se $\text{MCD}(x, y) = 1$

Osservazione 13. Data una soluzione non primitiva $\text{MCD}(x, y) = k$ si ha $x = ku$, $y = kv$ e

$$N = x^2 - Dy^2 = k^2(u^2 - Dv^2)$$

quindi $k^2 \mid N$ e per trovare una tale soluzione si possono prima trovare tutte le soluzioni primitive per $N' = \frac{N}{k^2}$ al variare di tutti i divisori quadratici di N.

Definizione 9 (Equivalenza delle soluzioni primitive)**:** Due soluzioni primitive $\alpha_1 = x_1 + y_1\sqrt{D}$, $\alpha_2 = x_2 + y_2\sqrt{D}$ di $x^2 - Dy^2 = N$ sono dette equivalenti se il loro rapporto è una soluzione $u + v\sqrt{D}$ dell'equazione di Pell $x^2 - Dy^2 = 1$.

Lemma 24 (Condizione per l'Equivalenza)**:** Condizione necessaria e sufficiente affinché α_1 e α_2 siano equivalenti è

$$x_1x_2 - Dy_1y_2 \equiv 0 \,(\text{mod } Q_0), \qquad x_1y_2 - y_1x_2 \equiv 0 \,(\text{mod } Q_0)$$

dove $Q_0 = |N|$.

Dimostrazione. Razionalizzando la frazione $\frac{\alpha_1}{\alpha_2}$ e coniugando ciò che risulta, imponendo che soddisfi l'equazione di Pell risulta:

$$\alpha_1 \sim \alpha_2 \Leftrightarrow \left(\frac{x_1x_2 - Dy_1y_2}{N}\right)^2 - D\left(\frac{x_2y_1 - x_1y_2}{N}\right)^2 = 1$$

Quindi il rapporto tra due di esse dà sempre una "soluzione" alla Pell, anche se essa può non avere coefficienti interi. La condizione che i coefficienti siano interi si esprime con il sistema di congruenze di cui sopra. □

Osservazione 14. Da ciò che abbiamo appena notato segue anche che, a meno di equivalenza, vi sono un numero finito di soluzioni *fondamentali* alla Pell per $c = N$, e tutte le altre si ottengono moltiplicando queste per una soluzione della Pell con $c = 1$.

Per dimostrare ciò notiamo che se abbiamo due soluzioni (x_1, y_1) e (x_2, y_2) di $x^2 - Dy^2 = N$ sugli interi tali che $x_1 \equiv x_2 \,(\text{mod } Q_0)$ e $y_1 \equiv y_2 \,(\text{mod } Q_0)$ allora esse sono equivalenti, ovvero si ha

$$x_2x_1 - Dy_1y_2 \equiv x_1^2 - Dy_1^2 \equiv 0 \,(\text{mod } Q_0)$$

$$x_1y_2 - y_1x_2 \equiv x_1y_1 - y_1x_1 \equiv 0 \,(\mathrm{mod}\ Q_0)$$

quindi le classi di equivalenza sono al più N^2.

5.2.2 Sistema di riduzioni di Lagrange

In questo sezione esponiamo il sistema di riduzioni di Lagrange che fornisce un algoritmo per controllare se l'equazione di Pell ha soluzione per d, c fissati. Il lemma principale non viene dimostrato.

Abbiamo già osservato in 9 che, dovendo risolvere l'equazione $x^2 - Dy^2 = N$ se $N < \sqrt{D}$, allora la soluzione $\frac{x}{y}$ è una convergente di $\sqrt{D}$.

Il metodo di riduzione di Lagrange ci viene in aiuto qualora dovessimo risolvere l'equazione $x^2 - Dy^2 = N$ con $N \geq \sqrt{D}$. Consideriamo infatti il seguente lemma:

Lemma 25: Se $x, y > 0$ è una soluzione di $x^2 - Dy^2 = N$ allora $\exists X, Y, K$ tali che

$$0 \leq K \leq \frac{|N|}{2}, \qquad h = \frac{K^2 - D}{N} \in \mathbb{Z}, \qquad X^2 - DY^2 = h$$

Inoltre si ha

$$x = \left|\frac{KX + DY}{h}\right|,\ y = \left|\frac{KY + X}{h}\right| \quad \text{oppure} \quad x = \left|\frac{KX - DY}{h}\right|,\ y = \left|\frac{KY - X}{h}\right|$$

Spesso il lemma viene applicato ricorsivamente, ovvero $\forall K \in [0, \frac{|N|}{2}]$ tale che $h = \frac{K^2-D}{N} \in \mathbb{Z}$ si riapplica il metodo di Lagrange all'equazione $X^2 - DY^2 = h$, ovvero se $h^2 < D$ la si risolve, altrimenti se $h^2 > D$ si riapplica la riduzione.

Si prende poi una soluzione da ciascuna classe di quelle trovate e si percorrono all'indietro le riduzioni effettuate per trovare una soluzione all'equazione originale.

Non sempre si hanno soluzioni con N generico, come si può banalmente osservare per congruenza prendendo $D \equiv 1 \,(\mathrm{mod}\ 4)$ e $N \equiv 2 \,(\mathrm{mod}\ 4)$. Visto che $x^2, y^2 \equiv 0, 1 \,(\mathrm{mod}\ 4)$ si ottiene l'assurdo.

5.3 Collegamento con i divisori e le curve iperellittiche

Supponiamo che $d \in K[t]$ sia squarefree. Allora possiamo considerare la curva iperellittica

$$u^2 = d(t)$$

che è non singolare per il criterio Jacobiano. Dimostriamo allora il seguente collegamento tra i divisori e l'esistenza di soluzioni all'equazione di Pell:

Teorema 26: Sia $\mathcal{C}$ il completamento proiettivo della curva non singolare $u^2 = d(t)$, che ha due punti all'infinito: siano essi ∞_- e ∞_+. Esiste una soluzione all'equazione di Pell $x(t)^2 - d(t)y(t)^2 = 1$ se e solo se il divisore $\infty_- - \infty_+$ è di torsione, ovvero $\exists m \in \mathbb{N}$ tale che $m(\infty_- - \infty_+) = 0 \in \mathrm{Jac}\ \mathcal{C}$.

Dimostrazione. Se esiste una soluzione all'equazione di pell possiamo considerare $\phi_+ = x(t) + uy(t)$ e $\phi_- = x(t) - uy(t)$ nell'algebra di funzioni sulla curva $K[u, t]$. Si ha allora che

$\phi_+\phi_- = 1$ su tutta la parte affine e quindi il divisore divϕ_+ è supportato solo all'infinito (non potendo annullarsi né avere poli sulla parte affine) ed ha grado zero, visto che ϕ_+ è una funzione. Deve allora essere necessariamente della forma $m(\infty_- - \infty_+)$ per qualche $m \in \mathbb{N}$.

Viceversa supponiamo di avere che il divisore $\delta = \infty_- - \infty_+$ è di torsione. Essendo $m(\infty_- - \infty_+) = 0$ nella varietà Jacobiana, per definizione vi è una funzione $f \in K[x,u]$ tale che div$f = m\delta$. Ogni funzione f si scrive in modo unico come $x(t) + uy(t)$. Sia $\sigma : \mathcal{C} \to \mathcal{C}$ l'automorfismo definito da $(u,t) \mapsto (-u,t)$. L'azione di questo automorfismo sulla funzione f ci dà $g = x(t) - uy(t)$ ed otteniamo che $fg = x(t)^2 - d(t)y(t)^2$, ma il divisore di fg è supportato solo all'infinito perché f è una funzione regolare sulla parte affine e g è ottenuta da un'automorfismo che manda la parte affine in sè stessa. Allora fg non ha né zeri né poli al finito ed è quindi costante, risultando così essere una soluzione dell'equazione di Pell. □

Questo ad esempio potrebbe dare un metodo effettivo per calcolare se un dato polinomio è Pelliano (ovvero esiste una soluzione all'equazione di Pell): basterebbe vedere se il divisore δ sopra definito è un punto di torsione della Jacobiana della curva iperellittica. A tal proposito si può vedere [13], dove viene fornito un algoritmo per il calcolo nel caso in cui K sia finitamente generato sul campo base.

Tale algoritmo è basato sulla buona riduzione delle curve iperellittiche modulo primi. Esemplifichiamo il metodo per $d \in \mathbb{Z}[t]$: per ogni primo p che non divida il discriminante del polinomio $d(t)$, otteniamo per riduzione delle curve iperellittiche $\mathcal{C}_p$ definite su $\mathbb{F}_p$. Vale inoltre che se η è di ordine n nello Jac $\mathcal{C}$ e $p \nmid n$, si ha che η rimane di ordine n in Jac $\mathcal{C}_p$.

Allora, preso $\delta \in$ Jac $\mathcal{C}$ definito come sopra, si può calcolare l'ordine di δ nelle riduzioni $\mathcal{C}_p$ per due precisi primi p, q (non ci soffermiamo sul modo in cui sceglierli che è però il busillis dell'intero metodo) che non dividano il discriminante di $d(t)$.

Per controllare su $\mathbb{F}_p[t]$ l'ordine di un punto nella Jacobiana si può utilizzare il metodo delle frazioni continue (che sono sempre periodiche in $\mathbb{F}_p[t]$).

Da qui attraverso l'osservazione precedente (la riduzione modulo p preserva la parte di ordine coprimo con p) si può ricostruire l'ordine di δ in Jac $\mathcal{C}$ (ed osservare se esso è finito od infinito).

Capitolo 6

Specializzazione di Soluzioni Minime

6.1 Motivazione e Alcune Osservazioni

Dato $D \in \mathbb{Z}[t]$ pelliano su $\mathbb{Z}[t]$, ovvero tale che esista una soluzione $(P, Q) \in \mathbb{Z}[t]$, dalla valutazione di quest'ultima sull'intero t_0 si ottiene sicuramente una soluzione dell'equazione di Pell per l'intero $D(t_0)$. Viene naturale chiedersi quando essa sia la minima, e se si riesca a dire qualche altra cosa a proposito.

Questa parte è stata ispirata da alcuni confronti con il Prof. Umberto Zannier, che ringrazio per il tempo concessomi.

Esempio 3: Supponiamo di avere $D = t^2 + t = t(t+1)$. Abbiamo visto che le soluzioni si possono esplicitamente scrivere. In particolare otteniamo su $\mathbb{C}[t]$ si ha $P(t) = 2t + 1$ e $Q(t) = 2$.

Per questa possiamo mostrare che per $t \neq 0, -1$ (per i quali $D(t) \leq 1$), le soluzioni date valutando P e Q sono le minime su $\mathbb{Z}$, ad esempio in $t = 2$ abbiamo $D = 6$ e $P = 5$, $Q = 2$.

6.1.1 L'idea della Dimostrazione della minimalità

Spieghiamo qui la motivazione per la quale si dovrebbe pensare che le valutazioni di una soluzione polinomiale minima debbano dare luogo ad una soluzione minima sugli interi. Questo paragrafo è completamente informale.

Esempio 4: Utilizziamo come esempio $D = t^4 + 4t^2 + 6$. Con il metodo delle frazioni continue si può vedere che la soluzione minima polinomiale è $P = t^4 + 4t^2 + 5$ e $Q = t^2 + 2$. Inoltre lo sviluppo in frazione continua di $\sqrt{D}$ è $[t^2 + 2, \overline{t^2 + 2, 2t^2 + 4}]$.

Come sappiamo la soluzione polinomiale è data in frazione continua da $\frac{P}{Q} = [t^2 + 2, t^2 + 2]$. Notiamo che tutti i polinomi che compaiono nella frazione continua, se valutati, danno sempre numeri interi (maggiori di due), quindi andando a valutare i polinomi presenti si ottiene una frazione continua naturale che dà una soluzione della Pell intera.

Siccome in generale un multiplo della soluzione fondamentale si scrive come:

$$\frac{x_s}{y_s} = [a_0; a_1, \ldots, a_n, \underbrace{2a_0, a_1, \ldots, a_n, \ldots, 2a_0, a_1, \ldots, a_n}_{b \text{ volte}}]$$

se la specificazione fosse un multiplo della fondamentale si dovrebbe avere che $t_0^2 + 2 = 2(t_0^2 + 2)$ il che ovviamente non è possibile. Abbiamo quindi mostrato che in questo caso tutte le specificazioni sono minime: ad esempio per $t_0 = 1$ si ha $D = 11$ e $P = 10, Q = 3$ che si può verificare essere la minima.

Purtroppo in generale i polinomi a_i possono assumere valori negativi (ad esempio $D = t^4+4t^2+3$) o possono anche avere valori razionali seppur P e Q siano polinomi a coefficienti interi (ad esempio $D = t^4 + 6t^3 + 5t^2 - 8t + 2$ che vedremo in seguito) perciò si rende necessario aggiustare l'idea grezza di cui sopra.

6.1.2 $D(t)$ è un quadrato perfetto per un numero finito di valori

Lemma 27: L'insieme dei $t_0 \in \mathbb{Z}$ tali che $D(t_0) = a^2$ con $a \in \mathbb{Z}$ per D pelliano è finito.

Dimostrazione. Se $D(t) = a^2$ si ha $(P(t)-aQ(t))(P(t)+aQ(t)) = 1$ e quindi $P(t)^2 = 1$, le cui soluzioni sono in numero finito. Inoltre le si può separare in due casi:

- $D(t) = 0$
- $D(t) \neq 0 \Rightarrow Q(t) = 0$

Queste ultime sono estremamente rare poiché si devono annullare contemporaneamente i due polinomi $P^2(t) - 1$ e $Q(t)$. □

6.1.3 La soluzione specificata non è sempre la minima

Facendo un po' di prove con polinomi semplici si potrebbe pensare che la soluzione specificata coincida sempre con la minima. Purtroppo non è sempre così e ne forniamo un controesempio:

Esempio 5: Consideriamo $D(t) = t^4 + 6t^3 + 9t^2 + t + 3$, la cui soluzione possiamo verificare essere

$$P(t) = 2t^3 + 6t^2 + 1, \qquad Q(t) = 2t$$

visto che $\sqrt{D(t)} = [t^2 + 3t, \overline{2t, 2t^2 + 3t}]$.

Valutando in $t = -4$ otteniamo $D(-4) = 15$, $|P(-4)| = 31$, $|Q(-4)| = 8$, mentre la soluzione fondamentale per $d = 15$ è $p = 4$, $q = 1$. In particolare abbiamo trovato la soluzione che è il doppio della fondamentale.

D'altra parte la soluzione polinomiale $P(t), Q(t)$ è quella minimale per come l'abbiamo ottenuta dall'algoritmo delle frazioni continue.

6.1.4 Unicità della scrittura in frazioni continue naturali

Lemma 28: Supponiamo di avere due diverse scritture in frazioni continue naturali (finite) di uno stesso numero $\alpha \in \mathbb{R}^+$:

$$\alpha = [a_0; a_1, \ldots, a_n] = [b_0; b_1, \ldots, b_m]$$

con $a_0, b_0 \geq 0$ e per $i \geq 1 \quad a_i, b_i \geq 1$ e con $a_n, b_m \geq 2$ (infatti se $a_n = 1$ si ha $[\ldots, a_{n-1}, 1] = [\ldots, a_{n-1} + 1]$).

Allora si ha $n = m$ e $a_i = b_i \quad \forall i$.

Dimostrazione. Sappiamo che $a_0 + \frac{1}{[a_1;\ldots,a_n]} = b_0 + \frac{1}{[b_1;\ldots,b_m]}$ e ovviamente si ha $[a_1; \ldots, a_n] \geq 1$ con uguaglianza se e solo se $n = 1$ (ed analogamente per i b_i). Allora $\frac{1}{[a_1;\ldots,a_n]} < 1$ da cui, siccome $a_0, b_0 \in \mathbb{N}$ si ottiene

$$a_0 = b_0, \qquad [a_1; \ldots, a_n] = [b_1; \ldots, b_m]$$

Da qui per induzione si può concludere (tralasciamo i dettagli per brevità). □

6.2 Limitatezza dell'esponente

Data la soluzione fondamentale in $\mathbb{Z}[t]$ della Pell polinomiale, valutando in $t_0 \in \mathbb{Z}$ ci si può chiedere che multiplo della soluzione fondamentale negli interi per $D(t_0)$ si ottenga, ovvero quali n possano comparire come:

$$P(t_0) + \sqrt{D(t_0)}Q(t_0) = (x_0 + \sqrt{D(t_0)}y_0)^n \tag{6.1}$$

In questa sezione mostreremo che, a $D(t)$ fissato, n è limitato da una costante effettivamente calcolabile.

6.2.1 Notazioni ed Identità matriciali

Definiamo per comodità la seguente notazione per le matrici:

$$\boxed{a} := \begin{pmatrix} a & 1 \\ 1 & 0 \end{pmatrix}, \qquad J := \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

e notiamo inoltre le seguenti identità matriciali:

$$\begin{aligned} \boxed{a}\boxed{0}\boxed{b} &= \boxed{a+b} \\ \boxed{a}\boxed{-b} &= \boxed{a-1}\boxed{1}\boxed{b-1}J \\ J\boxed{c} &\sim \boxed{-c}J \end{aligned}$$

dove con $A \sim B$ intendiamo che

$$A\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \lambda B \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

per qualche $\lambda \in \mathbb{Z}$.

Inoltre valgono le seguenti due identità:

$$\lambda \begin{pmatrix} 1 & r \\ 0 & \lambda \end{pmatrix}\begin{pmatrix} p & q \\ q & 0 \end{pmatrix} = \begin{pmatrix} p+rq & \lambda q \\ \lambda q & 0 \end{pmatrix}\begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix} \tag{6.2}$$

$$\begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} p & q \\ q & 0 \end{pmatrix} = \begin{pmatrix} \lambda p & q \\ q & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix} \tag{6.3}$$

6.2.2 Corrispondenza matriciale per le frazioni continue

Data una matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ con $a, b, c, d \in R$ possiamo interpretarla come funzione razionale $\mathbb{P}^1_L \to \mathbb{P}^1_L$: $z \mapsto r(z) = \frac{az+b}{cz+d}$ e la moltiplicazione di matrici coincide con la composizione di funzioni.

Allora diciamo che una frazione $\frac{a}{c}$ è in corrispondenza con tutte le funzioni razionali r tali che $r(\infty) = \frac{a}{c}$, ovvero con le matrici $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ per b, d qualunque. Notiamo che, con le notazioni di sopra, se $A \sim B$, si ha $A(\infty) = B(\infty)$.

6.2.3 Riduzione da frazioni continue intere a naturali

Lemma 29: Supponiamo di avere una frazione continua α intera matriciale, ovvero

$$\alpha \leftrightarrow \boxed{a_0} \dots \boxed{a_n}$$

con $a_i \in \mathbb{Z}$. Allora possiamo ottenere

$$\alpha \leftrightarrow \boxed{b_0} \dots \boxed{b_m}$$

con $b_j \in \mathbb{N}$ per $j \neq 0$ e $b_0 \in \mathbb{Z}$, $m \leq n + 2s$ con s il numero di cambi di segno nella stringa $a_0, \dots, a_n$. (Dove lo zero ha il segno più comodo, ovvero non viene contato come cambio di segno). In particolare siccome $s \leq \left\lfloor \frac{n-2}{2} \right\rfloor < \frac{n}{2}$ si ha $m < 2n$.

Inoltre la dipendenza dei b_j dagli a_i è di somme e differenze (dipendenti anche da costanti), subordinate ad una decisione di natura semialgebrica sugli a_i ($a_i > 0$ oppure $a_i = 0$).

Dimostrazione. Utilizzando le equazioni notate sopra otteniamo le seguenti uguaglianze con $a, b, c \in \mathbb{Z}$:

R0 $JJ = I$

R1 $\boxed{x}\boxed{0}\boxed{y} = \boxed{x+y}$

R2 $\boxed{x}\boxed{-y} = \boxed{x-1}\boxed{1}\boxed{y-1}J$

S0 $(a_i \geq 0, \forall d, b_j)$

$$\boxed{d}\boxed{a_1}\dots\boxed{a_k}J\boxed{-a_0}\boxed{b_1}\dots\boxed{b_r} = \boxed{d}\boxed{a_1}\dots\boxed{a_k}\boxed{a_0}\boxed{-b_1}\dots\boxed{-b_r}J$$
$$\sim \boxed{d}\boxed{a_1}\dots\boxed{a_k}\boxed{a_0}\boxed{-b_1}\dots\boxed{-b_r}$$

Data la stringa $\boxed{a}\boxed{b}\boxed{c}$ in modo che $b < 0$ (ci si può sempre ridurre in questa forma usando R1 se abbiamo abbastanza termini, oppure abbiamo terminato la riscrittura) si distinguono i casi seguenti:

S1 $(b = -1, c \leq 1, (a \geq 1$ `oppure` a `iniziale`$))$ $\boxed{a}\boxed{-1}\boxed{c} \Rightarrow \boxed{a-1}\boxed{1-c}J$
Se $a = 1$ facciamo seguire questa regola da R1, se si può applicare.
Inoltre se $d < 0$ usiamo la trasformazione S0.
Se invece $d > 0$ si ottiene la riduzione:
$(c \neq 1)$ $\boxed{a}\boxed{-1}\boxed{c} \Rightarrow \boxed{a-1}\boxed{-c}\boxed{1}\boxed{d-1}$
Per $c = 1$ si ha invece $\boxed{a}\boxed{-1}\boxed{1}\boxed{d} \Rightarrow \boxed{a-d-1}J$.
Se a `iniziale` sono a posto, utilizzando o la regola S0, oppure se $e > 0$ (o ancora se $a - d - 1 > 0$), si ottiene
$\boxed{a-d-1}J\boxed{e} \Rightarrow \boxed{a-d-2}\boxed{1}\boxed{e-1}$
mentre se a non è iniziale e $a - d - 1 < 0$, si ha s davanti posivito e quindi
$\boxed{s}\boxed{a-d-1}J \Rightarrow \boxed{s-1}\boxed{1}\boxed{d-a}$

S2 $(b = -1, c \geq 2, (a \geq 1$ `oppure` a `iniziale`$))$ $\boxed{a}\boxed{-1}\boxed{c} \Rightarrow \boxed{a-2}\boxed{1}\boxed{c-2}$
Nel caso in cui $a = 1$ davanti abbiamo un termine s che è ≥ 1 oppure il termine iniziale. In ambo i casi si ottiene:
$\boxed{s}\boxed{1}\boxed{-1}\boxed{c} \Rightarrow \boxed{s+1-c}J$
Se s è il termine iniziale siamo a posto (facendo i casi come sopra). Nel caso invece in cui $s + 1 - c < 0$ otteniamo, con termine t davanti che
$\boxed{t}\boxed{s}\boxed{1}\boxed{-1}\boxed{c} \Rightarrow \boxed{t-1}\boxed{1}\boxed{c-s-2}$
dove ora tutti i termini sono non negativi.

S3 $(b \leq -2, (a \geq 1$ `oppure` a `iniziale`$))$ $\boxed{a}\boxed{b} \Rightarrow \boxed{a-1}\boxed{1}\boxed{-b-1}J$
Se $c < 0$ si applica la S0 per levarsi la J
Se $c > 0$ si ottiene invece:
$\boxed{a}\boxed{b}\boxed{c} \Rightarrow \boxed{a-1}\boxed{1}\boxed{-b-2}\boxed{1}\boxed{c-1}$

SZ $\boxed{a}\boxed{b}\boxed{0}\star \Rightarrow \boxed{a}\star$

SE Nel caso in cui si abbia con $b \leq -1, (a \geq 1$ `oppure` a `iniziale`$)$ la stringa
$\boxed{a}\boxed{b}\star \Rightarrow \boxed{a-1}\boxed{1}\boxed{-b-1}\star$

Notiamo inoltre che la quantità calcolata globalmente $Q =$ "numero di termini + numero di matrici $J + 2\times s(a_0, \ldots, a_n)$" è strettamente decrescente in tutte le riscritture precedenti. Inoltre si ha

- Al momento iniziale si ha $Q = n + 2s$
- Al momento finale si ha $Q = m$, visto che sono tutti positivi (ed il primo si conta sempre come positivo)

e si ottiene quindi la tesi considerando che le trasformazioni sopra esposte sono una partizione delle possibilità che si presentano sul primo termine negativo nella stringa a partire da sinistra.

Inoltre, come si può vedere dalle premesse delle trasformazioni, le condizioni sui coefficienti a_i sono tutte di tipo semialgebrico. □

La stima ottenuta è ottimale

Esempio 6: La stima ottenuta sopra è ottimale: si consideri infatti $\boxed{3}\boxed{-3}\boxed{3}\boxed{-3}\ldots\boxed{3}$.

Utilizzando la regola di riscrittura

$$\boxed{a}\boxed{-b}\boxed{c} \Rightarrow \boxed{a-1}\boxed{1}\boxed{b-2}\boxed{1}\boxed{c-1}$$

otteniamo prima $\boxed{2}\boxed{1}\boxed{1}\boxed{1}\boxed{2}\boxed{-3}\boxed{3}\boxed{-3}\ldots\boxed{3}$.

Riapplicandola alla nuova sequenza ottenuta a partire dal secondo 2 otteniamo

$$\boxed{2}\boxed{1}\boxed{1}\boxed{1}\boxed{1}\boxed{1}\boxed{1}\boxed{1}\boxed{2}\boxed{-3}\ldots\boxed{3}$$

Continuando in questo modo, se partivamo da una stringa lunga $2n+1$, ad ogni passo aggiungiamo 2 termini e togliamo un solo numero negativo. In totale, visto che abbiamo n numeri negativi, la stringa finale ottenuta è lunga $(2n+1)+2n = 4n+1 = 2(2n+1)-1$ e che realizza quindi il massimo.

6.2.4 Caso con i polinomi in $Z[t]$

Quando abbiamo una soluzione all'equazione polinomiale per $D(t)$, scrivendo $\frac{P(t)}{Q(t)}$ in frazioni continue si ha, utilizzando la corrispondenza geometrica:

$$\frac{P}{Q} \leftrightarrow \begin{pmatrix} P & Q \\ Q & 0 \end{pmatrix} = \boxed{A_0}\ldots\boxed{A_n}\begin{pmatrix} 1 & R \\ 0 & \Lambda \end{pmatrix} \sim \boxed{A_0}\ldots\boxed{A_n}$$

Ora, quando andiamo a specificare le soluzioni sappiamo di ottenere delle soluzioni alla Pell sugli interi per $D(t)$ e quindi vale che (a meno di cambiare segno alla x od alla y)

$$\frac{x}{y} \leftrightarrow \begin{pmatrix} x & y \\ y & 0 \end{pmatrix} \sim \boxed{A_0(t)}\ldots\boxed{A_n(t)}$$

dove tutti gli $A_i(t) \in \mathbb{Z}$

Allora possiamo operare le riduzioni date dal lemma precedente ed ottenere una frazione continua nella quale compaiono solamente numeri naturali.

6.2.5 Riduzione da frazioni continue razionali ad intere

Potrebbe ora venire in mente che tutti i quozienti a_i debbano essere in $\mathbb{Z}[t]$. Ciò purtroppo non è vero, come possiamo vedere dal seguente:

Esempio 7: Consideriamo $D = t^4+6t^3+5t^2-8t+2$. Si verifica allora utilizzando l'algoritmo risolutivo attraverso le funzioni continue che la soluzione minima è data da $P = t^4+10t^3+31t^2+24t-17$ e $Q = t^2+7t+12$, ma lo sviluppo in frazione continua di $\sqrt{D}$ ci dà:

$$\sqrt{D} = [t^2+3t-2, \quad \frac{1}{2}t+\frac{7}{4}, \quad \ldots]$$

Si rende quindi necessario avere un algoritmo per espandere anche le frazioni continue razionali in frazioni continue intere.

Ci riproponiamo di mimare la dimostrazione precedente per ridurre anche le frazioni continue espresse con numeri razionali in frazioni continue intere (da cui, applicando il procedimento precedente, ci si riconduce a quelle naturali). Saranno necessarie un paio di premesse sull'algoritmo Euclideo per l'MCD.

Velocità dell'algoritmo Euclideo per l'MCD

Teorema 30: Dati $a, b \in \mathbb{Z}$ indichiamo con $R(a,b)$ (dove $a \leq b$) il numero di divisioni euclidee necessarie per calcolare MCD(a,b). Sia F_n l'n-esimo numero di Fibonacci: $F_0 = F_1 = 1$, $F_{n+2} = F_{n+1} + F_n$.

Se $a \leq F_n$ oppure $b \leq F_n$ si ha che $R(a,b) \leq n$.

Dimostrazione. La dimostrazione procede per induzione:

Passo Base $n = 1$: allora si ha $a \leq 1$ e quindi in al più un passo (dividere b per a) si ottiene l'MCD

Passo Induttivo: se $F_n < a \leq F_{n+1}$ (e q è il quoziente dato dalla divisione euclidea) deve essere che $b - qa \leq F_n$. Se per assurdo avessimo $a > b - qa > F_n$ avremmo anche

$$a = (b - qa) + qa > F_n + qa > F_n + qF_n > F_{n-1}$$

(poiché $q \geq 1$) ed otteniamo l'assurdo. □

Frazioni continue ed algoritmo euclideo

Se abbiamo una matrice della forma $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ con $a, b, c, d \in \mathbb{Z}$ possiamo "decomporla" in un po' di matrici della forma $\boxed{s}$ utilizzando l'algoritmo euclideo sulla prima colonna:

Siano q, r quelli dati dalla divisione euclidea $a = qc + r$. Allora si ha la decomposizione

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} q & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} c & d \\ r & b - qd \end{pmatrix}$$

Continuando ad effettuare questa decomposizione sull'ultima matrice a destra, otteniamo la *forma fondamentale* della matrice di partenza:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \boxed{q_0}\,\boxed{q_1} \cdots \boxed{q_n} \begin{pmatrix} 1 & r \\ 0 & \lambda \end{pmatrix}$$

Inoltre, sfruttando questa identità matriciale sulla prima colonna dell'ultima matrice si ottiene che $\boxed{q_0}\,\boxed{q_1} \ldots \boxed{q_n}$ è la forma della frazione continua di $\frac{a}{c}$.

Algoritmo per la riduzione delle frazioni continue razionali

Teorema 31: Supponiamo di avere una frazione continua α razionale matriciale, ovvero

$$\alpha \leftrightarrow \boxed{a_0} \ldots \boxed{a_n}$$

con $a_i \in \mathbb{Q}$. Allora possiamo ottenere

$$\alpha \leftrightarrow \boxed{b_0} \ldots \boxed{b_m}$$

con $b_j \in \mathbb{Z}$ per $j \neq 0$ ed m limitato solo in funzione di n e dei denominatori degli α_i.

Dimostrazione. Come notato sopra la frazione continua rappresentata in notazione geometrica è invariante per moltiplicazioni per scalare. Allora se $a_i = \frac{p_i}{q_i}$ otteniamo che $\boxed{a_i} \sim \begin{pmatrix} p_i & q_i \\ q_i & 0 \end{pmatrix}$ ovvero

$$\alpha \leftrightarrow \begin{pmatrix} p_0 & q_0 \\ q_0 & 0 \end{pmatrix} \cdot \ldots \cdot \begin{pmatrix} p_n & q_n \\ q_n & 0 \end{pmatrix}$$

Utilizzando l'osservazione precedente scriviamo la prima matrice in forma fondamentale:

$$\begin{pmatrix} p_0 & q_0 \\ q_0 & 0 \end{pmatrix} = \boxed{c_{0,0}} \boxed{c_{0,1}} \cdots \boxed{c_{0,k(0)}} \begin{pmatrix} 1 & r_0 \\ 0 & \lambda_0 \end{pmatrix}$$

dove, a causa della conservazione del determinante, $\lambda_0 = q_0^2$ e per quanto osservato sulla velocità di convergenza dell'algoritmo euclideo si può stimare $k(0)$ in funzione di q_0.

Usando le equazioni di scambio 6.2 con $\begin{pmatrix} p_1 & q_1 \\ q_1 & 0 \end{pmatrix}$ definiamo:

$$p_1' = p_1 + r_0, \qquad q_1' = \lambda_0 q_1$$

ottenendo quindi

$$\alpha \leftrightarrow \boxed{c_{0,0}} \boxed{c_{0,1}} \cdots \boxed{c_{0,k(0)}} \begin{pmatrix} p_1' & q_1' \\ q_1' & 0 \end{pmatrix} \begin{pmatrix} \lambda_0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} p_2 & q_2 \\ q_2 & 0 \end{pmatrix} \cdots \begin{pmatrix} p_n & q_n \\ q_n & 0 \end{pmatrix}$$

Effettuando la medesima operazione sulla matrice $\begin{pmatrix} p_1' & q_1' \\ q_1' & 0 \end{pmatrix}$ otteniamo

$$\begin{pmatrix} p_1' & q_1' \\ q_1' & 0 \end{pmatrix} = \boxed{c_{1,0}} \boxed{c_{1,1}} \cdots \boxed{c_{1,k(1)}} \begin{pmatrix} 1 & r_1 \\ 0 & \lambda_1 \end{pmatrix}$$

dove, come notato precedentemente $\lambda_1 = q_1'^2$ e $k(1)$ può essere stimato solo in funzione di q_1'.

Si può continuare ad espandere e commutare le matrici fino ad arrivare ad esprimere α in funzione di soli polinomi a coefficienti interi:

$$\alpha \leftrightarrow \boxed{c_{0,0}} \cdots \boxed{c_{0,k(0)}} \boxed{c_{1,0}} \cdots \boxed{c_{1,k(1)}} \cdots \boxed{c_{n,0}} \cdots \boxed{c_{n,k(n)}}$$

□

6.2.6 Limitatezza dell'esponente

Avendo discusso del modo in cui dipende la lunghezza delle espansioni dai coefficienti della frazione continua razionale, possiamo passare a dimostrare la limitatezza dell'esponente nell'equazione 6.1.

Come nel caso della frazione continua con polinomi in $\mathbb{Z}[t]$, se abbiamo una frazione continua con i polinomi $\mathbb{Q}[t]$, ovvero

$$\alpha = [S_0(t), S_1(t), \ldots, S_n(t)]$$

con $S_i \in \mathbb{Q}[t]$, possiamo scrivere $S_i(t) = \frac{A_i(t)}{q_i}$ dove $q_i \in \mathbb{Z}$ e $A_i(t) \in \mathbb{Z}[t]$ ed applicare il lemma precedente per ottenere una frazione continua con polinomi in $\mathbb{Z}[t]$ di lunghezza k, limitata dai q_i e da n (che sono valori fissati).

Adesso possiamo applicare l'algoritmo per ridurre questa frazione continua con polinomi in $\mathbb{Z}[t]$ in frazioni continue con numeri naturali (a t_0 fissato), tutte di lunghezza limitata da $2k$.

Quindi abbiamo mostrato che al variare di $t_0 \in \mathbb{Z}$, $\frac{P(t_0)}{Q(t_0)}$ ha una frazione continua di lunghezza limitata. Sappiamo anche che se $P(t_0), Q(t_0)$ è l'm-esima potenza della soluzione minima per $D(t_0)$, si scrive

$$\frac{P(t_0)}{Q(t_0)} = [a_0, a_1, \ldots, a_n, \underbrace{2a_0, a_1, \ldots, a_n}_{\text{ripetuta } m-1 \text{ volte}}]$$

Allora, avendo mostrato che $\frac{P(t_0)}{Q(t_0)}$ ha frazione continua di lunghezza limitata, diciamo di lunghezza al più r, segue che $P(t_0), Q(t_0)$ non può essere l'm-esima potenza della soluzione minima per $m > r$, data l'unicità della scrittura in frazioni continua naturali.

6.3 Minimalità delle soluzioni

In questa sezione mostriamo che le soluzioni dell'equazione polinomiale, quando specificate, danno luogo a soluzioni minime per tutti i punti tranne un numero finito.

6.3.1 Definizioni dei polinomi di Chebycheff T_n e U_n

Definiamo i polinomi bivariati $S_n(x, y)$ e $R_n(x, y)$ nel seguente modo:

$$R_n(x, y) + \sqrt{d}S_n(x, y) = (x + \sqrt{d}y)^n$$

Essi sono quindi esplicitamente dati da:

$$R_n(x, y) = \sum_{\substack{k=0 \\ k=2h}}^{n} \binom{n}{k} (dy^2)^h x^{n-k}$$

$$S_n(x, y) = y \cdot \sum_{\substack{k=0 \\ k=2h+1}}^{n} \binom{n}{k} (dy^2)^h x^{n-k}$$

Se aggiungiamo come ipotesi che si abbia $x^2 - dy^2 = 1$ essi possono essere riscritti in una forma più semplice:

$$R_n(x, y) = \sum_{\substack{k=0 \\ k=2h}}^{n} \binom{n}{k} (x^2 - 1)^h x^{n-k} =: T_n(x)$$

$$S_n(x, y) = y \cdot \sum_{\substack{k=0 \\ k=2h+1}}^{n} \binom{n}{k} (x^2 - 1)^h x^{n-k} =: yU_n(x)$$

Una definizione alternativa di $T_n(x)$ e $U_n(x)$ utilizza le seguenti formule per ricorrenza:

$$T_{n+1}(x) = xT_n(x) + (x^2-1)U_n(x) \qquad T_1(x) = x$$
$$U_{n+1}(x) = xU_n(x) + T_n(x) \qquad U_1(x) = 1$$

Essi sono anche noti rispettivamente come polinomi di Chebycheff del primo e del secondo tipo.

Osservazione 15. Osserviamo che vale la formula $U_n(A)^2(1-A^2) = 1 - T_n(A)^2$: essa si può banalmente verificare per $n = 1$ e per induzione sugli n successivi utilizzando le formule per ricorsione.

6.3.2 Caratterizzazione della non-minimalità

Avere una soluzione $(t_0, P(t_0), Q(t_0))$ non minimale significa che esistono $n, x, y \in \mathbb{Z}$ tali che soddisfano le seguenti equazioni:

$$P(t_0) - T_n(x) = 0$$
$$Q(t_0) - U_n(x)y = 0$$
$$x^2 - D(t_0)y^2 - 1 = 0$$

Per convincersene, basta scrivere le varie definizioni. Inoltre possiamo chiedere n primo senza perdita di generalità: infatti se $n = pm$ con p primo ed una soluzione specificata è la n-esima potenza di un'altra soluzione intera, essa sarà in particolare anche p-esima potenza della m-esima potenza di quest'ultima.

6.3.3 Sezioni polinomiali della prima equazione

Richiamiamo il risultato principale di Bilu e Tichy [2], che useremo a breve:

Teorema 32: Siano $f(x), g(x) \in \mathbb{Q}[x]$ due polinomi non costanti. Consideriamo l'equazione

$$f(x) = g(y) \tag{6.4}$$

Allora se l'equazione 6.4 ha infinite soluzioni razionali con denominatore limitato deve essere $f = \phi \circ f_1 \circ \lambda$ e $g = \phi \circ g_1 \circ \mu$ dove $\lambda(x), \mu(x) \in \mathbb{Q}[x]$ sono polinomi lineari, $\phi(x) \in \mathbb{Q}[x]$ e $(f_1(x), g_1(x))$ è una coppia standard su $\mathbb{Q}$ tale che l'equazione $f_1(x) = g_1(y)$ ha infinite soluzioni razionali con denominatore limitato.

Le coppie standard sono le cinque seguenti, dove $a, b \neq 0$, $m, n > 0$ interi e $p(x)$ è un polinomio non zero.

K1 $(x^m, ax^r p(x)^m)$, dove $0 \leq r < m$, $\mathrm{MCD}(r, m) = 1$

K2 $(x^2, (ax^2 + b)p(x)^2)$

K3 $(D_m(x, a^n), D_n(x, a^m))$ con $\mathrm{MCD}(m, n) = 1$ dove $D_m(x, a)$ è l'm-esimo polinomio di Dickson, definito da:

$$D_m(z + a/z, a) = z^m + (a/z)^m$$

K4 $(a^{-m/2}D_m(x,a), -b^{-n/2}D_n(x,b))$ con $\mathrm{MCD}(m,n) = 2$

K5 $((ax^2-1)^3, 3x^4-4x^3)$

Teorema 33: Supponiamo che vi siano un numero infinito di soluzioni (t_k, x_k) intere a

$$P(t) = T_n(x)$$

con n primo.

Allora esiste un polinomio $A(t) \in \mathbb{Q}[t]$ tale che $P(t) = T_n(A(t))$.

Dimostrazione. Per il teorema di Bilu e Tichy [2] si ha che deve essere del tipo

$$P = \phi \circ f \circ \lambda$$
$$T_n = \phi \circ g \circ \mu$$

dove (f,g) sono una delle *coppie standard* ivi descritte e λ, μ sono due polinomi lineari. Ma T_n è indecomponibile poiché $\deg T_n = n$ che è primo.

Quindi, per quanto riguarda g, si deve verificare una delle seguenti alternative:

- $g = \sigma$ lineare e $\phi = T_n \circ (\mu^{-1} \circ \sigma^{-1})$, da cui $P = T_n \circ A$ come volevasi.
- $g = \sigma \circ T_n \circ \mu^{-1}$ e $\phi = \sigma^{-1}$, ma questa seconda possibilità può capitare solo con una coppia standard del primo tipo nel caso in cui $p(x) = \sigma \circ T_n \circ \mu^{-1}$, $m = 1$, $r = 0$, $a = 1$. In questo caso però si ha $f = x$ e quindi P dovrebbe essere un polinomio lineare, il che è escluso dal fatto che abbia grado pari.

□

6.3.4 Sufficienza della risoluzione della prima equazione

Supponiamo di avere una soluzione $A(t)$ a $T_n(A(t)) = P(t)$. Utilizzando la seconda equazione definiamo allora $y = \frac{Q(t)}{U_n(A(t))}$. Supponendo ora che valga $P(t)^2 - D(t)Q(t)^2 = 1$ mostriamo che vale anche $A(t)^2 - D(t)y^2 = 1$ (seppur non sappiamo se $y \in \mathbb{Q}[t]$, potrebbe stare in $\mathbb{Q}(t)$):

Supponiamo n dispari (il caso $n = 2$ viene trattato in seguito) e quindi $T_n(A) = A \cdot R_n(A)$. Allora si ha:

$$\begin{aligned} A^2 - Dy^2 &= \frac{P^2}{R_n(A)^2} - \frac{DQ^2}{U_n(A)^2} \\ &= \frac{U_n(A)^2P^2 - R_n(A)^2(P^2-1)}{U_n(A)^2R_n(A)^2} \\ &= \frac{P^2(U_n(A)^2 - R_n(A)^2) + R_n(A)^2}{U_n(A)^2R_n(A)^2} \\ &= \frac{A^2R_n(A)^2(U_n(A)^2 - R_n(A)^2) + R_n(A)^2}{U_n(A)^2R_n(A)^2} \\ &= \frac{1 + A^2(U_n(A)^2 - R_n(A)^2)}{U_n(A)^2} \end{aligned}$$

Come visto precedentemente vale però $U_n(A)^2(1-A^2) = 1 - T_n(A)^2$ e questo ci dà la tesi.

Nel caso $n = 2$ invece si ha: $2A^2 - 1 = P$, $y = \frac{Q}{2A}$ e quindi

$$\begin{aligned} A^2 - Dy^2 &= \frac{P+1}{2} - \frac{P^2-1}{2(P+1)} \\ &= \frac{P+1}{2} - \frac{P-1}{2} = 1 \end{aligned}$$

6.3.5 Minimalità delle soluzioni

Abbiamo quindi appena mostrato che, se abbiamo infinite soluzioni intere a $P(t) = T_n(x)$ e $P(t)^2 - D(t)Q(t)^2 = 1$, esiste un polinomio $A(t) \in \mathbb{Q}[t]$ tale che $P(t) = T_n(A(t))$ ed inoltre esiste $y \in \mathbb{Q}(t)$ tale che $Q(t) = yU_n(A(t))$ e $A(t)^2 - D(t)y^2 = 1$. Se riuscissimo a mostrare che $y \in \mathbb{Q}[t]$ avremmo trovato un'altra soluzione polinomiale alla Pell.

Osservazione 16. Se scriviamo $y = \frac{B(t)}{C(t)}$ ai minimi termini, abbiamo

$$D(t)\left(\frac{B(t)}{C(t)}\right)^2 = A(t)^2 - 1 \in \mathbb{Q}[t]$$

da cui otteniamo che $C(t)^2 \mid D(t)$.

Se supponiamo quindi che $D(t)$ sia squarefree, otteniamo che $y \in \mathbb{Q}[t]$.

Abbiamo quindi trovato una soluzione $A(t) + \sqrt{D(t)}B(t) \in \mathbb{Q}[t, \sqrt{D(t)}]$ che elevata alla n-esima potenza dà la soluzione originaria $P(t) + \sqrt{D(t)}Q(t)$.

Teorema 34: Se $D(t) \in \mathbb{Z}[t]$ è squarefree e la soluzione minima in $\mathbb{Q}[t]$ $P(t), Q(t)$ sta in $\mathbb{Z}[t]$, allora per tutti i $t_0 \in \mathbb{Z}$ tranne un numero finito, la soluzione $P(t_0), Q(t_0)$ è la soluzione minima in $\mathbb{Z}$ per $D(t_0)$.

Dimostrazione. Se la soluzione specificata fosse non minima per infiniti t_0 si avrebbe, per il Teorema 33 e per le osservazioni precedenti, l'esistenza di due polinomi $A(t), B(t) \in \mathbb{Q}[t]$ tali che $P(t) + \sqrt{D(t)}Q(t) = (A(t) + \sqrt{D(t)}B(t))^n$, ma ciò contraddice le nostre ipotesi. □

Bibliografia

[1] N. H. Abel. "Ueber die Integration der Differential-Formel $\frac{\varrho d.x}{\sqrt{R}}$, wenn R und ϱ ganze Functionen sind". In: *J. Reine Angew. Math.* 1 (1826), pp. 185–221.

[2] Yuri Bilu e Robert Tichy. "The Diophantine equation $f(x) = g(y)$". In: *Acta Arithmetica* 95.3 (2000), pp. 261–288.

[3] Lou van den Dries e Karsten Schmidt. "Bounds in the theory of polynomial rings over fields. A nonstandard approach". In: *Inventiones mathematicae* 76.1 (1984), pp. 77–91.

[4] Michael J. Jacobson Jr. e Hugh C. Williams. *Solving the Pell equation*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. Springer, New York, 2009, pp. xx+495. ISBN: 978-0-387-84922-5.

[5] A. Ya. Khinchin. *Continued fractions*. The University of Chicago Press, Chicago, Ill.-London, 1964, pp. xi+95.

[6] R. C. Mason. *Diophantine equations over function fields*. Vol. 96. London Mathematical Society Lecture Note Series. Cambridge University Press, Cambridge, 1984, pp. x+125.

[7] Keith Matthews. "The diophantine equation $X^2 - DY^2 = N$, $D > 0$". In: *Expositiones Mathematicae* 18.4 (2000), pp. 323–332.

[8] Olaf Merkert. "Reduction and specialization of hyperelliptic continued fractions". Tesi di dott. Scuola Normale Superiore, 2016. arXiv preprint arXiv:1706.048011.

[9] Richard A Mollin e Anitha Srinivasan. "A Note on the negative pell equation". In: *International Journal of Algebra* 4.19 (2010), pp. 919–922.

[10] Alfred J. van der Poorten e Xuan Chuong Tran. "Quasi-elliptic integrals and periodic continued fractions". In: *Monatsh. Math.* 131.2 (2000), pp. 155–169.

[11] Noah Snyder. "An alternate proof of Mason's theorem". In: *Elem. Math.* 55.3 (2000), pp. 93–94.

[12] W. W. Stothers. "Polynomial identities and Hauptmoduln". In: *Quart. J. Math. Oxford Ser. (2)* 32.127 (1981), pp. 349–370.

[13] Jing Yu. "On arithmetic of hyperelliptic curves". In: *manuscript marked Aspects of Mathematics, Hong Kong University* (1999), pp. 4–6.