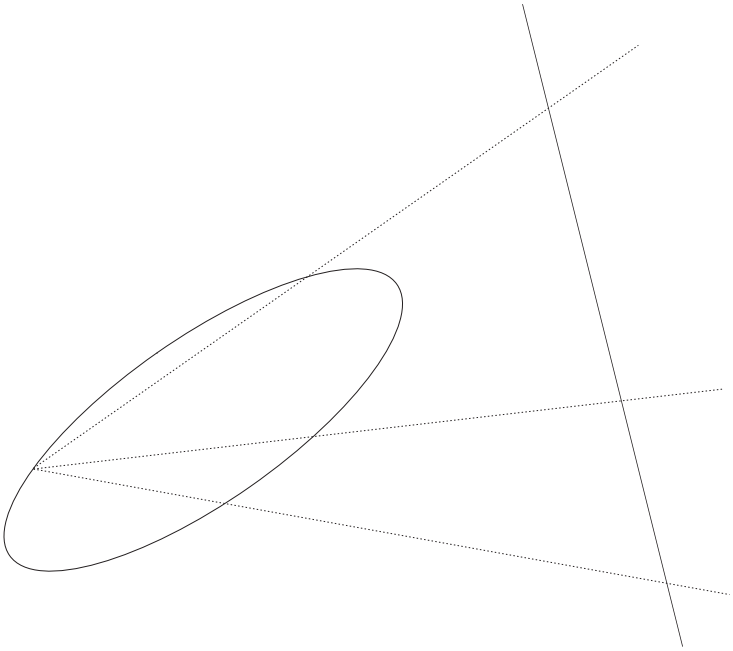
Aritmética em Retas e Cônicas



Rodrigo Gondim
Departamento de Matemática - UFRPE
rodrigo.gondim.neves@gmail.com
I Colóquio de Matemática do Nordeste
Sergipe, 28/02/2011 - 04/03/2011

Agradecimentos

Agradeço aos Deuses e aos Homens, de todas as civilizações, que inventaram, desenvolveram e exploraram os Números (Aritmética) e as Formas (Geometria)

Agradeço aos mestres Marcos Miguel, Antonio Carlos e Francesco Russo que, em diferentes momentos, me fizeram descobrir e me apaixonar pela Geometria e pela Aritmética.

Agradeço à Marco Mialaret Jr pelo empenho em estudar esses temas e por escrever, junto a mim, a monografia que deu origem à essas notas.

Agradeço aos amigos Gabriel Guedes, Gersonilo Silva, Hebe Cavalcanti e Tiago Duque pela leitura, pelas críticas e pelas sugestões.

Conteúdo

Introdução

As primeiras preocupações matemáticas do ser humano diziam respeito aos *números* e às *formas*. A Aritmética (do grego $\acute{\alpha}\rho\iota\theta\mu\acute{\varsigma}$ = número) foi concebida como o estudo dos números, em especial dos números inteiros. A geometria (do grego $\gamma\epsilon\omega\mu\tau\rho\acute{\iota}\alpha$) (em que $\gamma\epsilon\omega$ = terra e $\mu\tau\rho\acute{\iota}\alpha$ = medições, medida) é o estudo das formas - são casos especiais as retas, os polígonos e as cônicas. Vamos tentar explicar do que se trata *Aritmética em Retas e Cônicas*.
Há indícios históricos de que os problemas aritméticos ocupavam um lugar central entre o conhecimento matemático em várias civilizações antigas e medievais. Na Índia destacamos os matemáticos Baudhayana (800 AC), Apastamba (600 AC), Aryabhata e Bhaskara I (século VI), Brahmagupta (século VII), Mahaviara (século IX) e Bhaskara II (século XII). Na Grécia antiga (período helenístico): Pitágoras (500 AC), Arquimedes (300 AC), Euclides (300 AC), Diofanto (250). Na China Zhou Bi Suan (300 AC) e Ch'in Chiu-Shao. Citamos ainda os matemáticos Mulçumanos Ibn al-Haytham (séculos X-XI), Kamal al-Dim al-Farisi (séculos XIII-XIV). Todos esses matemáticos estavam intensamente preocupados com questões aritméticas e contribuíram para o desenvolvimento deste ramo da Matemática. Os principais livros desse período são *Os Elementos* de Euclides (300 AC), *Jiuzhang Suanshu* "Os nove capítulos da Arte Matemática"(autor chinês desconhecido-300 AC), *Aritmética* de Diofanto(250) e *Brahma Sphuta Siddhanta* de Brahmagupta(628) traduzido em Árabe(773) e Latin(1123).

A motivação inicial do estudo da aritmética estava centrada em problemas concretos com soluções inteira. Esses problemas ficaram famosos por serem (ou parecerem ser) *indeterminados*. Consideremos, por simplicidade, o seguinte exemplo:
(Problema proposto por Euler) Um grupo de homens e mulheres gastaram, em uma taverna, 1000 patacas. Cada homem gastou 19 patacas e cada mulher 13. Quantos eram os homens e quantas eram as mulheres?

Nos séculos XVII e XVIII nasce uma nova fase a partir dos trabalhos de Fermat, Euler, Gauss, Lagrange, Legendre. Gauss, em seu memorável *Disquisitiones Arithmeticae* (1801) escreve :
The celebrated work of Diophantus, dedicated to the problem of indeterminateness, contains many results which excite a more than ordinary regard for the ingenuity and proficiency of the author, because of their difficulty and the subtle devices he uses, especially if we consider the few tools that he had at hand for his work... The really profound discoveries are due to more recent authors like those men of immortal glory P. de Fermat, L. Euler, L. Lagrange, A. M. Legendre (and a few others).

Nessa nova fase foi o interesse intrínseco na Aritmética que impulsionou as novas descobertas.
De maneira parnasiana, a Aritmética passa a fazer "Matemática pela Matemática"e muitos de seus problemas, de enunciado simples e intrigantes, não têm conexão alguma com questões concretas. Alguns desses problemas duraram séculos até serem solucionados. Como, por exemplo o famoso
(Último Teorema de Fermat, 1637) É impossível para um cubo ser escrito como a soma de dois cubos ou uma quarta potência ser escrita como a soma de duas quartas potências ou, em geral, para qualquer número que é uma potência maior que a segunda ser escrito como a soma de duas potências com o mesmo expoente (demonstrado por Wiles, A. em 1995)

Nessa perspectiva, não nos propomos fazer contextualizações "artificiais", uma vez que as mesmas parecem ter efeito negativo, como dizia, hiperbolicamente, Maiakoviski em sua auto biografia "Eu mesmo": *Meus estudos: Mamãe e primas de vários graus ensinavam-me. A aritmética parecia-me inverossímil. Era preciso calcular peras e maçãs distribuídas a garotos. Contudo, eu sempre recebia e dava sem contar, pois no Cáucaso há frutas à vontade. Foi com gosto que aprendi a ler.*

Observamos que os problemas lineares indeterminados que foram estudados desde a antiguidade, algebricamente, são escritos da forma $ax + by = c$, com $a, b, c \in \mathbb{Z}$. Outros, mais difíceis, mas também milenares são $z^2 = x^2 + y^2$ (estudados na India, China e Grécia) e $x^2 - dy^2 = 1$ (estudados na Índia) ou $x^n + y^n = z^n$ (de Fermat). Eles possuem uma característica em comum: todos são "polinomiais"em várias variáveis.

Chamamos Equação Diofantina (em homenagem ao matemático Diofanto de Alexandria) uma equação polinomial (em várias variáveis) com coeficientes inteiros e da qual procuramos obter soluções inteiras (ou racionais).

O nosso enfoque é a ligação entre a Geometria e Aritmética, assim, para nós, uma Equação Diofantina em duas variáveis representa uma curva no plano. Soluções inteiras (ou racionais) da Equação Diofantina correspondem a pontos com coordenadas inteiras (ou racionais) na curva que chamaremos ponto inteiro (ou racional). O casamento entre a Aritmética e a Geometria foi muito bem sucedido e Fermat foi um dos pioneiros na utilização de métodos geométricos para resolver Equações Diofantinas.

Importantes questões aritméticas de caráter qualitativo só foram resolvidas graças à utilização de métodos geométricos cada vez mais sofisticados. Os avanços da Geometria Algébrica, a teoria de Esquemas e suas aplicações foram os principais responsáveis.

As principais questões qualitativas que podem ser feitas neste contexto são:

(i) Existência de pontos racionais;

(ii) Decisão entre a finitude ou infinitude do conjunto dos pontos racionais;

(iii) Existência de pontos inteiros;

(iv) Decisão entre a finitude ou infinitude do conjunto dos pontos inteiros.

No caso das curvas as questões (i), (ii) e (iv) já foram praticamente respondidas para grau menor ou igual a três. No caso geral apenas as questões (ii) e (iv) estão solucionadas, mas a matemática envolvida é muito avançada. Para citar, o importante teorema de Faltings, G. que implica, em particular, que toda curva plana lisa com coeficientes inteiros e grau maior que 3 possui somente um número finito de pontos racionais (1983).

Estaremos interessados em tratar alguns problemas aritméticos clássicos com uma abordagem geométrica elementar. Os problemas específicos que abordaremos envolvem a procura de soluções inteiras (ou racionais) em equações polinomiais a duas variáveis (com coeficientes inteiros) de graus um e dois. Geometricamente essas são as Retas e as Cônicas (daí o nome do minicurso, *Aritmética em Retas e Cônicas*).
Daremos um tratamento um pouco mais geométrico ao estudo das retas. No caso das cônicas, além de resultados gerais, faremos um estudo detalhado de dois problemas clássicos específicos: Inteiros que são Soma de dois Quadrados $x^2+y^2 = n$ e as equações de Pell-Fermat $x^2 - dy^2 = 1$. Para tratar esses problemas introduziremos a teoria de Minkowski no plano, de maneira auto contida. Daremos referências que seguem essa linha de raciocínio para aqueles que quiserem dar continuidade ao estudo da assim chamada *Geometria Aritmética*.

Capítulo 1

Aritmética em Retas

Consideremos os três problemas a seguir, como motivação inicial:

Exemplo 1.1. Uma criança afirma que suas 1000 bolinhas de gude puderam ser guardadas algumas em latas grandes com 65 bolinhas cada uma e outras em latas menores com 26 bolinhas cada uma, de modo que todas as latas estavam completas. Reflita sobre a afirmação da criança, procurando descobrir a quantidade de latas grandes e a quantidade de latas pequenas.

Exemplo 1.2. Um pai, no começo do ano letivo, teve que comprar livros e cadernos para seus filhos. Cada livro custou $R\$50,00$ e cada caderno $R\$17,00$. Sabendo que o pai gastou $R\$570,00$ determine a quantidade de livros e cadernos comprados.

Exemplo 1.3. Um general decide dividir seu batalhão em colunas de 31 soldados e percebe que sobram 4 então tentou dividi-los em colunas de 50 soldados cada, desta vez sobrou um único soldado. Determine o número de soldados deste batalhão sabendo que tal número é menor que 1500.

Os três são problemas similares nos quais os métodos algébricos elementares para resolvê-los, via equações, se depara com um grande inconveniente:

UMA EQUAÇÃO, DUAS INCÓGNITAS

Esse tipo de problema comumente é dito ser *INDETERMINADO*. Entretanto, o tipo de incógnita implícita nestes problemas é inteira (e positiva)...

Geometricamente, uma equação linear em duas incógnitas representa uma reta no plano, ou seja, todas as soluções de um problema desses, no conjunto dos números reais, correspondem a uma reta no plano. Se procuramos soluções inteiras destas equações lineares, então estamos buscando pontos desta reta que estão situados sobre o reticulado inteiro do plano, $\mathbb{Z}^2 \subset \mathbb{R}^2$.

Problema 1.1. *Mostre que o problema do exemplo 1.1 não possui solução inteira, ou seja, a criança não falava a verdade.*

1.1 Pontos Inteiros em Retas

Considere uma reta $\ell = \{(x, y) \in \mathbb{R}^2 | bx + cy = a\}$ com coeficientes inteiros $a, b, c \in \mathbb{Z}$. Supondo que essa reta possui algum ponto $(x_0, y_0) \in \mathbb{Z}^2$ com coordenadas inteiras, então podemos concluir que a mesma possui uma infinidade de pontos inteiros.

Com efeito, se $v = (c, -b)$ é um vetor diretor desta reta, então os pontos $(x_k, y_k) = (x_0 + kc, y_0 - kb)$ com $k \in \mathbb{Z}$ são pontos inteiros da reta. O problema consiste em determinar se esses são *todos* os pontos inteiros da reta. É claro que se $mdc(b, c) = d \neq 1$, então devemos ter que $d|a$ pois $d|b$ e $d|c$ implica que $d|bx_0 + cy_0 = a$. Em outras palavras, as equações em que $d \not| a$ não possuem solução inteira.

No caso em que $d|a$ podemos escrever $b = d\beta$, $c = d\gamma$ e $a = d\alpha$ e encontrar uma nova equação (simplificada)

$$\beta x + \gamma y = \alpha$$

que é uma equação da mesma reta ℓ que possui um ponto inteiro $(x_0, y_0) \in \mathbb{Z}^2$. Um vetor diretor simplificado para esta reta é $w = (\gamma, -\beta)$. Redefinindo $(x_k, y_k) = (x_0 + k\gamma, y_0 - k\beta)$ em que $k \in \mathbb{Z}$, concluímos que todos os pontos (x_k, y_k) são pontos inteiros da reta e

a verdade é que estes são *todos* os pontos inteiros da reta ℓ.

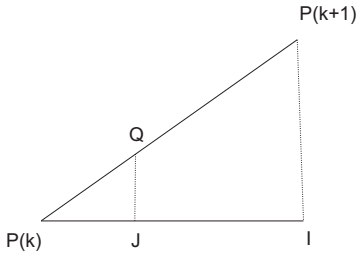
Podemos pensar nesse processo como uma simplificação do vetor diretor inteiro da reta ℓ. Com efeito, o vetor diretor de uma reta está definido a menos de múltiplo escalar (não nulo) e, portanto, a menos de sentido, existe um único vetor diretor cujas coordenadas são inteiros coprimos (supondo a existência de um vetor diretor inteiro). Tal vetor será chamado o vetor inteiro irredutível.

Proposição 1.2. *Seja $l = \{(x, y) \in \mathbb{R}^2 | bx + cy = a\}$ uma reta com coeficientes inteiros $a, b, c \in \mathbb{Z}^2$. Suponhamos que a reta possua um ponto inteiro $(x_0, y_0) \in \mathbb{Z}^2$. Seja $w = (-\gamma, \beta)$ o vetor diretor inteiro e irredutível da reta l, então todos os pontos inteiros da reta são:*

$$(x_k, y_k) = (x_0 + k\gamma, y_0 - k\beta)$$

em que $k \in \mathbb{Z}$.

Demonstração: É claro que todos esses pontos são pontos inteiros da reta, pois, a menos de sinal, $w = \frac{v}{d}$ em que $v = (c, -b)$ e $d = mdc(b, c)$. Logo, $bx_k + cy_k = bx_0 + cy_0 + b\frac{c}{d} - c\frac{b}{d} = bx_0 + cy_0 = 0$, pois $(x_0, y_0) \in l$. Só nos resta mostrar que esses são todos os pontos inteiros da reta.
Suponhamos, por absurdo, que exista um outro ponto inteiro em l, $Q = (x^*, y^*) \in \ell \cap \mathbb{Z}^2$. Digamos que, no sentido de nossa parametrização, esse ponto esteja entre os pontos inteiros P_k e P_{k+1}. Isso dá origem a dois triângulos retângulos semelhantes com hipotenusa sobre a reta e catetos nas direções horizontal e vertical.



O cateto horizontal do menor triângulo tem comprimento inteiro $h < |\gamma|$ e o cateto vertical do menor triângulo tem comprimento inteiro $s < |\beta|$. Por outro lado, pela proporcionalidade, temos que

$$\frac{h}{s} = \frac{|\gamma|}{|\beta|}$$

e isso é um absurdo pois os valores h, s, $|\gamma|$ e $|\beta|$ são todos inteiros positivos e a fração $\frac{|\gamma|}{|\beta|}$ é irredutível pois, por hipótese, $mdc(\gamma, \beta) = 1$.
□

Problema 1.3. *Encontre uma solução particular do segundo problema, exemplo 1.2, e use a proposição anterior para encontrar todas as soluções inteiras e positivas do mesmo. (Após parametrizar você deve impor que $x > 0$ e $y > 0$.)*

1.1.1 O Algoritmo Estendido de Euclides

Primeiramente, vamos relembrar o algoritmo de Euclides para o cálculo de mdc de dois números inteiros. Sejam b e c dois números inteiros e suponhamos que $b \geq c > 0$. Dividindo b por c obtemos quociente q e resto r, $r < c$ satisfazendo:

$$b = cq + r.$$

É fácil mostrar que:

$$mdc(b, c) = mdc(c, r).$$

Mais precisamente, o conjunto dos divisores comuns de b e c coincide com o conjunto dos divisores comuns de c e r. De fato, se $d|b$ e $d|c$, então $d|r = b - cq$ e, reciprocamente, se $d|c$ e $d|r$, então $d|b = cq + r$. Se procedermos, iteradamente, com este método, como os restos vão diminuindo, após um número finito de iterações obtemos resto igual a zero e então o algoritmo pára. Quando o resto for igual a zero, obtemos uma relação de divisibilidade $e|f$ neste caso, $mdc(e, f) = e$ é igual ao último divisor obtido.

Teorema 1.4. Algoritmo de Euclides

Sejam $b > c > 0$ dois números inteiros. Existe um algoritmo efetivo para encontrar o $mdc(b, c)$ a partir de (um número finito de) sucessivas divisões com resto.

Demonstração: *Algoritmo*

Sejam $r_{-1} = b$, $r_0 = c$ e façamos as divisões sucessivas de r_k por r_{k+1} em que $k = -1, 0, ..., n$ (em que n é o passo em que o resto é zero).

$$r_k = r_{k+1}q_{k+1} + r_{k+2}.$$

Pelo anteriormente exposto $mdc(r_k, r_{r_{k+1}}) = mdc(r_{k+1}, r_{r_{k+2}})$. O algoritmo funciona bem pois sempre temos $r_k > r_{k+1}$ e, como todos são não negativos($r_k \geq 0$), então $r_{n+1} = 0$ para algum n. Nesse caso $r_{n-1} = r_n q_n$ e, portanto, $mdc(r_{n-1}, r_n) = r_n$.

Logo,

$$mdc(b, c) = ... = mdc(r_k, r_{k+1}) = ... = mdc(r_{n-1}, r_n) = r_n$$

□

O algoritmo estendido de Euclides é um método interativo de obter, para cada um dos restos encontrados no algoritmo de Euclides, uma expressão do tipo:

$$r_k = bx_k + cy_k$$

na qual x_k e y_k são inteiros.

O algoritmo estendido de Euclides nos permite demonstrar o seguinte teorema conhecido por *Lema de Bèzout*:

Teorema 1.5. Algoritmo Estendido de Euclides - Lema de Bèzout

Sejam b e c números inteiros, $b > c > 0$, e seja $d = mdc(b, c)$. Então existem inteiros x e y tais que:

$$bx + cy = d.$$

Além disso, os inteiros x e y podem ser efetivamente calculados a partir de um algoritmo finito.

Demonstração: Com as mesmas notações estabelecidas no teorema anterior, vamos proceder interativamente para determinar x_k e y_k de modo que $bx_k + cy_k = r_k$.

Claramente, se $k = -1$, então podemos escolher $x_{-1} = 1$ e $y_{-1} = 0$ pois $1.b + 0.c = b$. Se $k = 0$ podemos escolher $x_0 = 0$ e $y_0 = 1$, pois $0.b + 1.c = c$. Agora, suponhamos que $k \geq 1$ e suponhamos que já conhecemos x_{k-2}, y_{k-2} tais que $bx_{k-2} + cy_{k-2} = r_{k-2}$ e também conhecemos x_{k-1} e y_{k-1} tais que $bx_{k-1} + cy_{k-1} = r_{k-1}$.

Como r_k é obtido a partir da divisão de r_{k-2} por r_{k-1}, temos:

$$r_{k-2} = r_{k-1}q_{k-1} + r_k$$

substituindo as expressões anteriores para r_{k-2} e r_{k-1}, obtemos:

$$r_k = b(x_{k-2} - q_{k-1}x_{k-1}) + c(y_{k-2} - q_{k-1}y_{k-1}).$$

Assim, podemos escolher $x_k = x_{k-2} - q_{k-1}x_{k-1}$ e $y_k = y_{k-2} - q_{k-1}y_{k-1}$.

O algoritmo vai parar quando $r_{n+1} = 0$, e no passo anterior, temos, pelo teorema 1.4, que

$$d = mdc(b, c) = r_n = bx_n + cy_n.$$

□

1.1.2 Equações Diofantinas Lineares

O teorema que exibiremos a seguir responde de uma vez por todas às indagações sobre a aritmética das retas e sua demonstração é uma leitura das seções anteriores.

Teorema 1.6 (Euclides). *Sejam a, b e c números inteiros e defina $d = mdc(b, c)$. Considere a equação diofantina linear:*

$$bx + cy = a$$

Então temos:

1. *(i) Se* $d \nmid a$*, então a equação dada não tem solução inteira;*

2. *(ii) Se* $d \mid a$*, então a equação dada tem infinitas soluções inteiras e, além disso, o conjunto de pontos inteiros pode ser parametrizado a partir de um ponto particular. Precisamente, fazendo* $a = d\alpha$*,* $b = d\beta$ *e* $c = d\gamma$ *a equação fica da forma:*

$$\beta x + \gamma y = \alpha$$

e dada uma solução particular (x_0, y_0) *nos inteiros, então todas as outras soluções inteiras da equação são da forma:*

$$\begin{cases} x &= x_0 + \gamma k \\ y &= y_0 - \beta k \end{cases} \quad k \in \mathbb{Z}$$

Demonstração: A primeira parte é trivial. A segunda parte segue da proposição 1.2 e do teorema 1.5. De fato, do teorema 1.5 concluímos que a reta possui um ponto inteiro (um ponto inteiro pode ser determinado pelo algoritmo estendido de Euclides) e, pela proposição 1.2, sabemos que se uma reta com coeficientes inteiros possui um ponto com coordenadas inteiras, então possui uma infinidade de pontos inteiros parametrizados a partir do ponto particular com múltiplos inteiros do vetor diretor irredutível. □

1.1.3 Áreas VS Pontos Inteiros em Regiões Poligonais

Existem muitas formas de calcular a área de um polígono no plano. Vamos apresentar duas delas: a primeira via determinantes e a segunda, será o teorema de Pick, que fornece uma maneira combinatória de calcular área de polígonos no plano cujos vértices pertencem ao "reticulado padrão" $\mathbb{Z}^2 \subset \mathbb{R}^2$, ou seja, seus vértices possuem coordenadas inteiras. A fórmula de Pick fornece uma relação entre a área de um polígono simples e o número de pontos inteiros em seu interior e em sua fronteira.

Área de um Paralelogramo

Um paralelogramo no plano pode ser definido a partir de dois vetores $v, w \in \mathbb{R}^2$. Seus vértices serão 0, v, w e $v + w$. Se considerarmos estes vetores como vetores espaciais (com a última coordenada nula), estamos fazendo a identificação $\mathbb{R}^2 \cong \{(x, y, z) \in \mathbb{R}^3 | z = 0\}$. Agora, podemos usar o produto vetorial do $\mathbb{R}^3$ para calcular a área do paralelogramo, verifica-se facilmente que

$$v \times w = \det(v, w)\hat{k} = (0, 0, \det(v, w)).$$

Ou seja, a área de um paralelogramo no plano gerado pelos vetores v e w é

$$A = ||v \times w|| = |\det(v, w)|.$$

Nessa notação, $\det(v, w)$ é o determinante da matriz quadrada de ordem 2 cujas linhas são, respectivamente, as coordenadas de v e w.

Sejam $v \in \mathbb{Z}^2 \subset \mathbb{R}^2$, $v = (a, b)$ e $d = mdc(a, b)$. Vamos dar uma interpretação geométrica do número d em termos de área de parelogramos no plano. Considere o seguinte conjunto $\Delta(v) = \{\det(v, w) | w \in \mathbb{Z}^2\}$, então temos o seguinte

Teorema 1.7. *Sejam $v \in \mathbb{Z}^2 \subset \mathbb{R}^2$, $v = (a, b)$, $d = mdc(a, b)$ e $\Delta(v) = \{\det(v, w) | w \in \mathbb{Z}^2\}$. Então:*

$$\Delta(v) = d\mathbb{Z} = \{dm | m \in \mathbb{Z}\}.$$

Ou seja, o mdc entre as coordenadas do vetor v representa a menor área de um paralelogramo com vértices inteiros tendo v como um dos lados.

Demonstração: Ora,

$$\Delta(v) = \{\det(v, w) | w \in \mathbb{Z}^2\} = \{ax + by | x, y \in \mathbb{Z}\} = d\mathbb{Z}$$

O resultado segue diretamente do lema de Bèzout, 1.5. □

O Teorema de Pick

O teorema de Pick fornece uma maneira combinatória de calcular a área de um polígono simples com vértices inteiros. A fórmula de Pick envolve o número de pontos inteiros na fronteira do polígono e o número de pontos inteiros no interior do polígono. No presente contexto a fórmula de Pick pode ser interpretada como uma forma de encontrar o número de pontos inteiros no interior do polígono.

Definição 1.8. *Um polígono plano é dito ser simples se não possuir "furos"e se suas arestas só se intersectarem nos vértices. Um polígono simples pode ser côncavo ou convexo.*

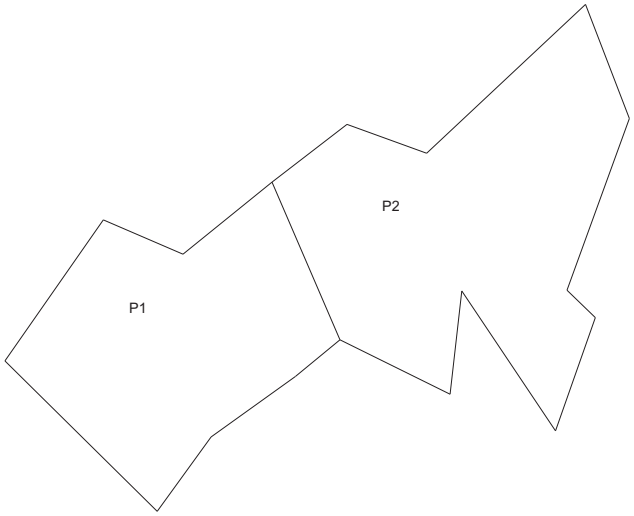
Teorema 1.9. (Pick) *Sejam $\mathcal{P} \subset \mathbb{R}^2$ um polígono simples cujos vértices pertencem a $\mathbb{Z}^2$ (pontos do plano com coordenadas inteiras). Defina F o número de pontos inteiros na fronteira de $\mathcal{P}$ (vértices e arestas) e I o número de pontos inteiros no interior de $\mathcal{P}$. Então a área do polígono $\mathcal{P}$ é*

$$A(\mathcal{P}) = \frac{1}{2}F + I - 1$$

Demonstração: Defina o número de Pick de um polígono simples $\mathcal{P}$ com vértices inteiros por:

$$Pick(\mathcal{P}) = \frac{1}{2}F + I - 1.$$

Se dois polígonos simples possuem uma aresta de mesmo módulo e direção, então podemos obter, a partir deles, um novo polígono identificando essa aresta e deletando-a, o polígono assim obtido é o que chamaremos a concatenação dos polígonos iniciais $\mathcal{P} = \mathcal{P}_1 \bigoplus \mathcal{P}_2$.



Vamos mostrar que o número de Pick é aditivo por concatenação. Sejam $\mathcal{P}_i$, $i = 1, 2$ dois polígonos simples com vértices inteiros e com uma aresta de mesmo módulo e direção. E sejam F_i e I_i, respectivamente, o número de pontos inteiros na fronteira e no interior do polígono $\mathcal{P}_i$, $i = 1, 2$. Digamos que o segmento comum, na concatenação possui $k + 2$ pontos inteiros.

O número de pontos inteiros no interior da concatenação é

$$I = I_1 + I_2 + k$$

pois, após a concatenação, os k vértices (não terminais) da aresta deletada vão pertencer ao interior do polígono $\mathcal{P} = \mathcal{P}_1 \oplus \mathcal{P}_2$.

O número de pontos inteiros na fronteira da concatenação é

$$F = F_1 + F_2 - 2(k + 2) + 2$$

pois somando os pontos de fronteira de $\mathcal{P}_1$ e $\mathcal{P}_2$ e subtraindo duas vezes os pontos inteiros da aresta deletada só faltam os terminais da aresta deletada para completar os pontos inteiros na fronteira de $\mathcal{P}$.

Calculando o número de Pick de $\mathcal{P}$, temos:

$$Pick(\mathcal{P}) = \frac{1}{2}(F_1+F_2-2(k+2)+2)+I_1+I_2+k = Pick(\mathcal{P}_1)+Pick(\mathcal{P}_2).$$

Agora note que todo polígono simples no plano pode ser subdividido em triângulos de modo que o vértice de cada triângulo seja algum vértice do polígono. Assim, todo polígono com vértices inteiros pode ser subdividido em triângulos com vértices inteiros. Pelo resultado de aditividade por concatenação, podemos nos reduzir ao caso de triângulos com vértices inteiros para provar o teorema de Pick.

Todo triângulo com vértices inteiros pode ser inscrito em um retângulo horizontal com vértices inteiros. Assim podemos nos reduzir aos triângulos retângulos horizontais ou melhor, aos próprios retângulos horizontais.

Todo triângulo horizontal é formado por concatenação de quadrados 1×1. Assim, se verificamos a fórmula de Pick em quadrados 1×1, então vale o teorema de Pick em geral

Para um quadrado 1×1 temos: $A = 1$, $F = 4$, $I = 0$, e efetivamente,

$$A = 1 = \frac{1}{2}.4 + 0 - 1 = \frac{1}{2}F + I - 1.$$

□

Observação 1.10. Ver, por exemplo, Lages Lima, E. [9]. A seção intitulada *Como calcular a área de um polígono se você sabe contar*

Como havíamos dito queremos olhar para a fórmula de Pick de outra forma:

$$I = A - \frac{1}{2}F + 1$$

E sabemos que a área pode ser calculada de várias formas. Gostaríamos de terminar essa seção mostrando como calcular F de maneira instantânea.

Proposição 1.11. *Seja $\mathcal{P} = A_0A_1A_2...A_{n-1}A_n$, com $A_n = A_0$, um polígono simples no plano com vértices inteiros, isto é, $A_i \in \mathbb{Z}$. Defina $v_i = \overrightarrow{A_{i-1}A_i} = (a_i, b_i)$ e $d_i = mdc(a_i, b_i)$. Então o número de pontos inteiros na fronteira de $\mathcal{P}$ é igual a $F = \sum_{i=1}^{n} d_i$.*

Demonstração: Primeiramente lembramos que, a partir do argumento utilizado na proposição 1.2, um segmento de reta $\overline{PQ}$ com $P, Q \in \mathbb{Z}^2$ tal que $v = \overrightarrow{PQ} = (a, b)$ com $mdc(a, b) = 1$ não possui ponto inteiro no seu interior. Ou seja, os únicos pontos inteiros no segmento $\overline{PQ}$ são seus extremos, P e Q. Verifique na figura 1.1!!!

Sejam $P, Q \in \mathbb{Z}^2 \subset \mathbb{R}^2$ e $v = \overrightarrow{PQ} = (a, b)$ dois vértices consecutivos do polígono, então o número de pontos inteiros na aresta $\overline{PQ}$ é igual a $d + 1$ em que $d = mdc(a, b)$. Com efeito, basta dividir o segmento $\overline{PQ}$ em d segmentos cujo vetor que o representa tenha coordenadas inteiros coprimos.

Para concluir note que cada aresta $\overline{A_{i-1}A_i}$ do polígono vai possuir, em seu interior(sem contar os vértices), $d_i - 1$ pontos inteiros. Logo o número de pontos inteiros na fronteira do polígono será

$$\sum_{i=1}^{n} d_i - n + n$$

$-n$ corresponde a -1 para cada aresta e $+n$ corresponde aos vértices do polígono. □

1.1.4 Problemas

1. Você possui muitos palitos com $6cm$ e $7cm$ de comprimento. Qual o número mínimo de palitos que você precisa utilizar para fazer uma fila de palitos com comprimento total de 2 metros?

2. (Problema proposto por Mahavira, 850) 5 pilhas de frutas mais duas frutas foram divididas (igualmente) entre 9 viajantes; seis pilhas mais quatro foram divididas por 8; quatro pilhas mais 1 foram divididas por 7. Determine o menor número possível de frutas em cada pilha.

3. (Problema proposto por Bhaskara 1; Século VI) Encontre o menor número natural que deixa resto 1 quando dividido por 2,3,4,5,6 mas é exatamente divisível por 7.

4. (Proposto por Euler) Uma pessoa comprou cavalos e bois. Foram pagos 31 escudos por cavalo e 20 escudos por boi e sabe-se que todos os cavalos custaram 7 escudos a mais do que todos os bois.Quantos cavalos e quantos bois foram comprados?

5. (Problema do século XVI) Um total de 41 pessoas entre homens, mulheres e crianças foram a um banquete e juntos gastaram 40 patacas. Cada homem pagou 4 patacas, cada mulher 3 patacas e cada criança um terço de pataca. Quantos homens, quantas mulheres e quantas crianças havia no banquete?

6. Na Russia a moeda se chama rublo. Existem notas de 1, 3 e 5 rublos. Mostre que não é possível pagar 25 rublos com exatamente 10 notas com os valores citados.

7. Prove que toda quantia inteira maior ou igual a $R\$4,00$ pode ser paga utilizando notas de $R\$2,00$ e $R\$5,00$.

8. Sejam $a, b, c \in \mathbb{Z}$ números inteiros positivos e suponha que $a \geq bc$. Mostre que a equação

$$bx + cy = a$$

possui solução inteira não negativa $(m, n) \in \mathbb{Z}$, $m, n \geq 0$

9. Dado $v \in \mathbb{Z}^2 \subset \mathbb{R}^2$, $v = (a, b)$, com $mdc(a, b) = 1$, mostre que existe $w = (c, d) \in \mathbb{Z}^2 \subset \mathbb{R}^2$ tal que o paralelogramo gerado por v e w não possui ponto inteiro em seu interior.

10. Prove que todo polígono simples com vértices inteiros possui área cujo dobro é um número inteiro.

11. Mostre que um triângulo plano com vértices inteiros tem área mínima $A = \frac{1}{2}$ se, e somente se, o triângulo não possui ponto inteiro no seu interior. Mostre que de fato essa é a área mínima.

12. Mostre que no plano existem triângulos com vértices inteiros de área mínima $A = \frac{1}{2}$ com perímetro arbitrariamente grande.

13. Um agricultor possui um terreno poligonal e deseja plantar pés de milho em seu interior. Suponhamos que, após uma escolha de eixos coordenados os vértices do polígono e os pés de milho vão corresponder a pontos com coordenadas inteiras. Determinar o número de pés de milho que podem ser plantados supondo que os vértices do polígono são: $A = (0, 0,)$, $B = (8, 0)$, $C = (15, 10)$, $D = (12, 20)$, $E = (10, 15)$ e $F = (0, 10)$.

Capítulo 2

Aritmética em Cônicas

2.1 Introdução

Uma curva algébrica plana em $\mathbb{R}^2$ é dada como o conjunto solução de uma equação polinomial em duas variáveis.

$$\mathcal{C} = \{(x, y) \in \mathbb{R}^2 \mid f(x, y) = 0\}$$

E consideramos que esse conjunto é não vazio e não consiste de um número finito de pontos. Chamamos cônica a uma curva algébrica plana definida por um polinômio de grau dois. Não estaremos interessados no caso em que o polinômio seja redutível, pois, nesse caso seu conjunto de zeros será um par de retas. É muito conhecido, da geometria analítica básica, que as únicas cônicas irredutíveis são: parábola, hipérbole e elipse.

As formas canônicas, com coeficientes inteiros, das cônicas são as seguintes:

1. **Parábolas**

$$\mathcal{C} = \{(x, y) \in \mathbb{R}^2 \mid ax^2 + by = 0\}$$

$a, b \in \mathbb{Z}, a \neq 0 \ e \ b \neq 0$

2. **Hipérboles**

$$\mathcal{C} = \{(x, y) \in \mathbb{R}^2 \mid ax^2 - by^2 = c\}$$

$a, b, c \in \mathbb{Z}, a, b, c > 0$

3. **Elipses**

$$\mathcal{C} = \{(x, y) \in \mathbb{R}^2 \mid ax^2 + by^2 = c\}$$

$a, b, c \in \mathbb{Z}, a, b, c > 0$

4. **Círculos** São elipses especiais para as quais $a = b$

$$\mathcal{C} = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = r^2\}$$

Observação 2.1. Quando estamos interessados no conjunto de pontos racionais de uma cônica, então podemos efetivamente nos reduzir à sua forma canônica. De fato, o método de diagonalização de uma forma quadrática funciona sobre qualquer corpo de característica diferente de 2.

Entretanto não é verdade o mesmo sobre os inteiros! Estaremos, de fato, fazendo uma simplificação ao supor uma cônica na forma canônica para tratar de questões sobre seu conjunto de pontos inteiros.

2.2 O Método das Tangentes e das Secantes de Fermat

Nessa Seção vamos apresentar o Método das Tangentes e das Secantes de Fermat. Tal método foi utilizado por Fermat em cônicas e cúbicas, por motivações aritméticas. Explicaremos o método somente para as cônicas.

Podemos sumarizar o Método das Secantes e das Tangentes de Fermat da seguinte forma:

Sejam $\mathcal{C}$ uma cônica, $P \in \mathcal{C}$ e ℓ uma reta que não passa por $P = P_1$. Seja $\tilde{\ell}$ a reta paralela a ℓ passando por P e $\mathcal{C} \cap \tilde{\ell} = \{P_1, P_2\}$

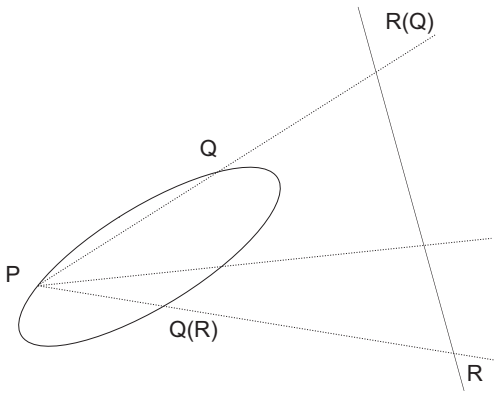
(não necessariamente distintos). Defina, ainda, $\tilde{P} = T_P C \cap \ell$. Então a função:

$$\begin{array}{rccc} \phi: & \mathcal{C} \setminus \{P_1, P_2\} & \longrightarrow & \ell \setminus \{\tilde{P}\} \\ & Q & \mapsto & \overline{PQ} \cap \ell = R(Q) \end{array}$$

está bem definida e é inversível. De fato, para cada $Q \in \mathcal{C} \setminus \{P_1, P_2\}$, a reta $\overline{PQ}$ não é paralela a ℓ (a única paralela a ℓ passando por P é $\overline{P_1 P_2} = \tilde{\ell}$). Assim, a intersecção $\overline{PQ} \cap \ell = R(Q)$ define um ponto, $R(Q)$.
A inversa de ϕ é a seguinte função:

$$\begin{array}{rccc} \phi^{-1}: & \ell \setminus \{\tilde{P}\} & \longrightarrow & \mathcal{C} \setminus \{P_1, P_2\} \\ & R & \mapsto & \overline{PR} \cap \mathcal{C} = Q(R) \end{array}$$

de fato, para que ϕ^{-1} pudesse ser definida, tivemos que excluir $\tilde{P}$, caso contrário a reta $\overline{\tilde{P}P}$ seria tangente a C.



Teorema 2.2. *Sejam $C \in \mathbb{R}^2$ uma cônica com coeficientes racionais e $P \in C$ um ponto racional. Considere $\ell \in \mathbb{R}^2$ uma reta com coeficientes racionais paralela a $T_P C$ (não coincidente). Então a função*

$$\begin{array}{rccc} \phi: & \mathcal{C} \setminus \{P\} & \longrightarrow & \ell \\ & Q & \mapsto & \overline{PQ} \cap \ell = R(Q) \end{array}$$

está bem definida, é bijetiva e leva pontos racionais de C em pontos racionais de ℓ.

Demonstração: Só nos resta mostrar que pontos racionais são levados em pontos racionais.
Para encontrar a interseção de uma reta e uma cônica devemos resolver o sistema de equações em duas variáveis consistindo de uma equação linear e uma quadrática. Substituindo $y = mx+n$, a equação da reta, na equação da cônica, o sistema fica reduzido a uma equação do segundo grau em uma única variável. Digamos que a equação do segundo grau seja

$$ax^2 + bx + c = 0.$$

Como a reta e a cônica possuem coeficientes racionais, então $a, b, c \in \mathbb{Q}$. Sabemos que as raízes da equação do segundo grau fornece abcissas dos pontos de interseção. Se temos uma raiz racional, então a outra raiz será também racional (pois o produto das raízes é igual a $\frac{c}{a}$, que é racional). Verifique!!!

Corolário 2.3. *Seja $C \subset \mathbb{R}^2$ uma cônica com coeficientes racionais. Suponhamos que o conjunto dos pontos racionais de C seja não vazio, então a cônica C possui uma infinidade de pontos racionais. Além disso, o conjunto dos pontos racionaqis de C pode ser parametrizado a partir do conhecimento de um ponto racional de C.*

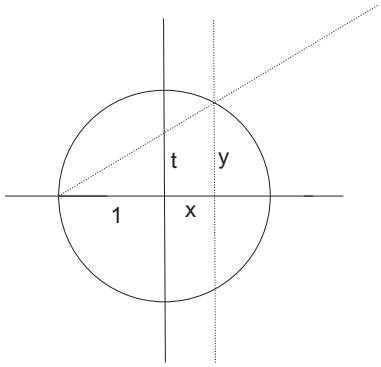
Demonstração: Basta utilizar o Método das Tangentes e das Secantes de Fermat. □

Consideremos, agora dois exemplos: o primeiro fornece uma parametrização do círculo unitário (sem usar senos e cossenos).

Exemplo 2.1. Seja $\mathcal{C} = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$. Usaremos o ponto $P_0 = (-1, 0)$ e a reta $\ell = \overline{OY} = \{(0, t) | t \in \mathbb{R}\}$ para aplicar o Método de Fermat:

$$\begin{array}{rccc} \phi: & \mathcal{C} \setminus \{P\} & \longrightarrow & \ell \\ & Q & \mapsto & \overline{PQ} \cap \ell \\ & (x, y) & \mapsto & (0, \frac{y}{x+1}) \end{array}$$

A parametrização que queremos é ϕ^{-1} e para encontrá-la, basta fazer $\frac{y}{x+1} = t \Rightarrow y = t(x + 1)$ e lembrar que $x^2 + y^2 = 1$. Daí segue que $(t^2 + 1)x^2 + (2t^2)x + (t^2 - 1) = 0$ e como $x_0 = -1$ é uma raiz



desta equação (correspondente ao ponto $P_0 = (-1, 0)$) e o produto das raízes é $\frac{t^2-1}{t^2+1}$ obtemos para a outra raiz $x_t = \frac{1-t^2}{1+t^2}$ substituindo temos $y_t = \frac{2t}{1+t^2}$. Logo,

$$\begin{array}{rccc} \phi^{-1}: & \ell & \longrightarrow & \mathcal{C} \setminus \{P\} \\ & Q & \mapsto & \overline{PQ} \cap \mathcal{C} \\ & (0,t) & \mapsto & (\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}) \end{array}$$

Exemplo 2.2. Considere agora o círculo $\mathcal{C} = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 = 2\}$. Um ponto racional nesse círculo é $(1, 1)$. Considere a família de retas passando por esse ponto: $\ell_t : y - 1 = t(x - 1)$, se $t \in \mathbb{Q}$, então cada uma dessas retas é secante ao círculo em um outro ponto racional P_t. Mostre que todos os outros pontos racionais desse círculo são:

$$(\frac{t^2 - 2t - 1}{t^2 + 1}, \frac{-t^2 - 2t + 1}{t^2 + 1}).$$

Teorema 2.4. *Seja $\mathcal{C} \subset \mathbb{R}^2$ a cônica de equação $ax^2 + by^2 = c$ com $a, b, c \in \mathbb{Q}$. Se $P_0 = (x_0, y_0)$ é um ponto racional de $\mathcal{C}$, então todos os outros pontos racionais de $\mathcal{C}$ são da forma*

$$(\frac{bt^2x_0 - 2bty_0 - ax_0}{bt^2 + a}, \frac{-bt^2y_0 - 2atx_0 + ay_0}{bt^2 + a})$$

em que $t \in \mathbb{Q}$, $bt^2 + a \neq 0$, exceto $(x_0, -y_0)$. Ou seja, o conjunto dos pontos racionais de C pode ser parametrizado a partir de P_0.

Demonstração: Considere a família de retas passando por P_0, $\ell_t : y - y_0 = t(x - x_0)$. Para cada parâmetro racional $t \in \mathbb{Q}$, a reta ℓ_t possui coeficientes racionais e, portanto, essa reta será secante à cônica $\mathcal{C}$ em um outro ponto racional $P_t = (x_t, y_t)$ (argumento de Fermat). Substituindo a equação da reta $y = y_0 + t(x - x_0)$ na equação da cônica obtemos uma equação do segundo grau:

$$(a + bt^2)x^2 + (\star)x + x_0(bt^2x_0 - 2bty_0 - ax_0) = 0$$

o produto das raízes dessa equação é:

$$x_0x_t = \frac{x_0(bt^2x_0 - 2bty_0 - ax_0)}{a + bt^2},$$

daí seguem as expressões de x_t e y_t. □

Observação 2.5. A partir do teorema acima podemos obter uma parametrização de qualquer cônica utilizando, apenas, funções racionais. O ponto central em nossa abordagem era encontrar soluções racionais, por isso a necessidade de o ponto inicial ser racional. Caso o objetivo seja encontrar uma parametrização utilizando funções racionais, então o ponto inicial pode ser tomado como um ponto qualquer $(x_0, y_0) \in \mathbb{R}^2$. Observamos ainda que as parábolas são triviais uma vez que sua equação fornece imediatamente uma parametrização das mesmas.

2.3 Homogeneização e Deshomogeneização: Curvas Projetivas

Um polinômio em mais de uma variável é dito ser homogêneo se todos os seus monômios são de mesmo grau, digamos n. Sua propriedade fundamental é que $F(\lambda P) = \lambda^n F(P)$, assim sendo, se $\lambda \neq 0$ então: $F(P) = 0 \Leftrightarrow F(\lambda P) = 0$.

Suponhamos, agora, que o polinômio F possua coeficientes racionais e vamos nos concentrar em procurar soluções racionais. Pelo exposto, quando estamos procurando soluções de um polinômio homogêneo, podemos nos reduzir a procurar soluções que não são múltiplos,

logo, na classe de soluções existe uma única solução inteira sem fatores comuns(a menos de sinal).

É bastante natural considerar a seguinte relação de equivalência em $\mathbb{R}^3 \setminus \{0\}$: $v \equiv w \Leftrightarrow v = \lambda w$ com $\lambda \neq 0$.

Definição 2.6. O Plano projetivo $\mathbb{P}^2(\mathbb{R})$ é definido como o conjunto quociente do $\mathbb{R}^3$ pela relação de equivalência de múltiplos não nulos. Se a classe de $v = (X, Y, Z)$ possuir $z \neq 0$, então $(X, Y, Z) \equiv (x, y, 1)$ caso contrário dizemos que o mesmo representa um ponto no infinito, assim

$$\mathbb{P}^2(\mathbb{R}) = \mathbb{R}^2 \cup \ell_\infty$$

em que ℓ_∞ representa a "reta no infinito"(dos pontos no infinito) que corresponde a $\mathbb{P}^1(\mathbb{R})$.

Se $f(x, y)$ é um polinômio em duas variáveis de grau (máximo) n podemos a partir de f obter um polinômio homogêneo $F(X, Y, Z)$ fazendo $x = \frac{X}{Z}$ e $y = \frac{Y}{Z}$ e cancelando denominadores. Observe que como estamos interessados em fazer $F(X, Y, Z) = 0$ é permitido cancelar denominadores.

Definição 2.7. Seja $f(x, y)$ um polinômio em duas variáveis com coeficientes reais e $C \subset \mathbb{R}^2$ a curva associada. O polinômio homogêneo F, associado a f é chamado homogeneização de f e $\overline{C} \subset \mathbb{P}^2$ a curva projetiva associada.

Proposição 2.8. *Seja $f(x, y)$ um polinômio com coeficientes racionais e $F(X, Y, Z)$ a homogeneização de f. Existe uma bijeção entre as soluções racionais de f e as soluções inteiras, sem fator comum e com $Z \neq 0$ de F (a menos de sinal).*

Demonstração: Observe que a cada solução racional em $f(x, y) = 0$ obtemos uma única solução em inteiros coprimos de $F(X, Y, Z) = 0$ (a menos de sinal). Basta multiplicar $(x, y, 1)$ pelo mmc das frações irredutíveis que determinam x, y.

Reciprocamente, se $F(X, Y, Z)$ é um polinômio homogêneo em três variáveis, podemos a partir de F obter um polinômio $f(x, y)$ fazendo $X = xZ$ e $Y = yZ$ e cancelando Z^n. Observamos que a a

cada solução em inteiros de $F(X,Y,Z)=0$, com $Z\neq 0$, obtemos uma solução racional de $f(x,y)=0$. □

Exemplo 2.3. Ternas Pitagóricas

Um problema milenar, conhecido por "ternas pitagóricas", consiste em encontrar soluções em inteiros para a equação de Pitágoras, $a^2=b^2+c^2$, isto é, encontrar triângulos retângulos com lados inteiros. Inicialmente vamos considerar o caso em que a, b e c são coprimos, isto é, $mdc(a,b,c)=1$ que implica, por sua vez, $mdc(a,b)=mdc(a,c)=mdc(b,c)=1$. Para encontrar a solução geral basta multiplicar por um inteiro arbitrário.

Dividindo a equação por a^2, quando $a\neq 0$, obtemos $\frac{b}{a}^2+\frac{c}{a}^2=1$, isto é, $x^2+y^2=1$. Desta feita, procuramos, agora, pontos racionais no círculo unitário. Sabemos que todos os pontos racionais no círculo unitário, exceto $(-1,0)$, podem ser parametrizados da forma $(\frac{1-t^2}{1+t^2},\frac{2t}{1+t^2})$ conforme visto no exemplo 2.1.

Suponhamos, agora que $t=\frac{m}{n}$, com $mdc(m,n)=1$. Substituindo, obtemos:

$$\frac{b}{a}=\frac{n^2-m^2}{n^2+m^2},\ \frac{c}{a}=\frac{2mn}{n^2+m^2}.$$

Como $mdc(a,b)=mdc(a,c)=1$, e b e c não podem ser ambos ímpares, temos $mdc(n^2-m^2,n^2+m^2)=1$ (m,n tem paridades distintas) nesse caso

$$a=m^2+n^2, b=n^2-m^2, c=2mn.$$

Alguns exemplos de ternas pitagóricas são $(3,4,5),(5,12,13),(8,15,17),(7,24,25),...$

Observação 2.9. *Poderia também ocorrer de m,n serem ambos ímpares. Nesse caso $mdc(n^2-m^2,n^2+m^2)=2$ e daí*

$$a=\frac{m^2+n^2}{2}, b=\frac{n^2-m^2}{2}, c=mn.$$

Fazendo $m+n = 2u$ e $n-m = 2v$ (que é sempre possível pois a soma e a diferença de dois ímpares é sempre par), obtemos

$$a = u^2 + v^2, b = 2uv, c = u^2 - v^2.$$

2.4 O Princípio Local-Global para as Cônicas

O princípio Local-Global, de Hasse-Minkowski, fornece uma caracterização das formas quadráticas (homogêneas) com coeficientes racionais (inteiros) que possuem algum ponto racional (não nulo). Após tal caracterização é possível resolver o problema algoritmicamente, isto é, dada uma forma quadrática é possível, após um número finito de passos, descobrir quando a mesma possui algum ponto racional. A formulação geral do princípio exige a definição de números p-ádicos e não é nosso objetivo. Vamos primeiramente mostrar, com um exemplo, que existem cônicas que não possuem pontos racionais e enunciar, posteriormente uma versão do princípio local-global para as cônicas.

Analisemos o seguinte exemplo, próximo ao exemplo das ternas pitagóricas. Vamos ver que uma ligeira modificação nos coeficientes de uma cônica pode trazer resultados catastróficos (mas interessantes). Nesse caso a cônica em questão não possui ponto racional.

Exemplo 2.4. Analisemos o conjunto dos pontos racionais do círculo

$$\mathcal{C} = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 = 3\}$$

ou equivalentemente, as soluções não nulas, em inteiros, da equação homogeneizada:

$$X^2 + Y^2 = 3Z^2,$$

podemos supor que se houver solução em inteiros $mdc(X, Y, Z) = 1$. Fazendo a divisão euclidiana (com resto) de cada um deles por 3, obtemos

$$\begin{cases} X &= 3\tilde{X} + a \\ Y &= 3\tilde{Y} + b \\ Z &= 3\tilde{Z} + c. \end{cases}$$

Substituindo na equação original notamos que $a^2 + b^2 - 3c^2$ deve ser múltiplo de 3 mas como $a, b, c \in \{0, 1, 2\}$ chegamos a conclusão (após alguns testes) que $a = b = 0$ e, portanto, X *e* Y deveriam ser múltiplos de 3 mas isso obrigaria que $3Z^2$ fosse múltiplo de 9 e isto só seria possível se Z fosse também múltiplo de 3 e isto é um absurdo uma vez que supomos que $mdc(X, Y, Z) = 1$.

Esta contradição foi proveniente da nossa (falsa) hipótese de existência de alguma solução em inteiros não nulos de $X^2 + Y^2 = 3Z^2$, assim podemos concluir que tal equação não possui solução não nula em inteiros e conseqüentemente, a nossa curva $\mathcal{C}$ não possui ponto racional.

Observação 2.10. A estratégia utilizada para mostrar que a equação não possui solução inteira não nula é chamada *Descida Infinita de Fermat*. Essa estratégia é baseada no princípio da Boa Ordenação. A filosofia geral da "descida infinita de Fermat"é a seguinte: se supomos que existem soluções inteiras positivas, podemos escolher uma minimal (em um sentido a determinar) e, a partir desta encontrar outra "menor", então não existe solução inteira. No nosso caso a minimalidade dizia respeito a não existir fator comum entre as coordenadas da solução.

Em geral temos o seguinte teorema de Hasse-Minkowski que determina quando uma cônica possui ponto racional.

Teorema 2.11 (**Hasse-Minkowski).** *Seja* $C \in \mathbb{R}^2$ *uma cônica com coeficientes racionais e* $\overline{C} \subset \mathbb{P}^2$ *a curva projetiva associada (com coeficientes inteiros). Uma condição necessária e suficiente para a existência de um ponto racional em* C *é que a sua equação homogeneizada possua solução módulo* p^e *para todo natural primo* p *e para cada* $e > 0$.

2.5 Uma Visão Geral sobre a Aritmética das Cônicas

2.5.1 Elipses

(i) Existência de ponto(s) racional(is).

Pelo Princípio Local-Global para as Cônicas, Teorema 2.11 o problema é algorítmico. Daremos uma descrição completa para as circunferências, independente do Teorema de Hasse-Minkowski, no capítulo intitulado Soma de Dois Quadrados.

(ii) Decisão entre a finitude ou infinitude do conjunto dos pontos racionais.

Pelo teorema 2.3, se a cônica C possui um ponto racional, então possui infinitos. Além disso, se co-nhecemos um deles, então é possível parametrizar todos, pelo teorema 2.4.

(iii) Existência de ponto(s) inteiro(s)

Muito difícil, em geral, faremos uma análise detalhada do caso do círculo no capítulo intitulado Soma de Dois Quadrados.

(iv) Decisão entre a finitude ou infinitude do conjunto dos pontos inteiros.

Esse conjunto é sempre finito, pela compacidade das elipses. Verifique!!!

2.5.2 Parábolas

(i) Existência de ponto(s) racional(is);

Toda parábola possui uma infinidade de pontos racionais.

(ii) Decisão entre a finitude ou infinitude do conjunto dos pontos racionais;

Toda parábola possui uma infinidade de pontos racionais.

(iii) Existência de ponto(s) inteiro(s);

Existem parábolas que possuem e outras que não possuem pontos inteiros.

(iv) Decisão entre a finitude ou infinitude do conjunto dos pontos inteiros.

A análise de exemplos mostrará que não há um padrão geral.

2.5.3 Hipérboles

(i) Existência de ponto(s) racional(is).

Pelo Princípio Local-Global para as Cônicas, Teorema 2.11 o problema é algorítmo. Faremos uma análise detalhada de uma classe especial de Hipérboles denominadas de Pell-Fermat na seção homônima.

(ii) Decisão entre a finitude ou infinitude do conjunto dos pontos racionais.

Pelo teorema 2.3, se a cônica C possui um ponto racional, então possui infinitos. Além disso, se conhecemos um deles, então é possível parametrizar todos, pelo teorema 2.4.

(iii) Existência de ponto(s) inteiro(s).

Muito difícil, em geral, faremos uma análise de um caso clássico chamadas Equações de Pell-Fermat no capítulo homônimo.

(iv) Decisão entre a finitude ou infinitude do conjunto dos pontos inteiros.

A análise de exemplos mostrará que não há um padrão geral, isto é, existem hipérboles que possuem infinitos pontos inteiros e outras que possuem somente um número finito de pontos inteiros.

2.6 Problemas

1. Seja $m \in \mathbb{Z}$ um inteiro positivo. Mostre que as parábolas $y^2 = mx$ possuem uma infinidade de pontos inteiros.

2. Mostre que as parábolas a seguir não possuem pontos inteiros

   (a) $y^2 = 4x + 2$

   (b) $y^2 = 4x + 3$

   (c) $y^2 = 3x + 2$

3. Encontre uma infinidade de pontos inteiros nas hipérboles

   (a) $x^2 - 3y^2 = 1$

   (b) $x^2 - 5y^2 = 1$

   (c) $x^2 - 7y^2 = 1$

4. Seja $n \in \mathbb{Z}$ um inteiro positivo. Mostre que a curva

$$\frac{1}{x} + \frac{1}{y} = \frac{1}{n}$$

   é uma hipérbole. Mostre que tal hipérbole tem um número finito de pontos inteiros qualquer que seja n.

5. Encontre todos os pontos inteiros e positivos da curva

$$\frac{1}{x} + \frac{1}{y} = \frac{1}{n}$$

   nos casos em que $n = p$ é primo, $n = p^2$ é o quadrado de um primo e $n = pq$ é o produto de dois primos.

6. Mostre que se uma cônica com coeficientes racionais possuir uma reta tangente com coeficientes racionais, então o ponto de tangência é racional.

7. Mostre que a equação

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

   não possui solução inteira não nula.

8. Determine todos os pares de inteiros (x, y) tais que

$$9xy - x^2 - 8y^2 = 2005.$$

9. Mostre que não existem pontos racionais na cônica $x^2 + xy + y^2 = 2$.

10. Encontre todos os pontos racionais a cônica $x^2 + xy + y^2 = 1$.

11. Mostre que as soluções inteiras coprimos da equação

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

são $x = \pm(u^2 - 2v^2)$, $y = 2uv$ e $z = u^2 + 2v^2$, com u, v inteiros primos entre si.

Capítulo 3

Reticulados no Plano

3.1 Reticulados e seus Domínios Fundamentais

A teoria de reticulados é bem desenvolvida e um leitor com uma maior familiaridade a teoria de grupos (abelianos) apreciará a leitura de Stewart [1]. Nosso objetivo é apresentar a teoria em um caso muito especial, reticulados no plano; pois, nesse caso, muitas construções e demonstrações se simplificam. Em particular utilizaremos o mínimo possível de álgebra linear.

Definição 3.1. Sejam $v_1, v_2 \subset \mathbb{R}^2$ dois vetores não múltiplos. Um reticulado (de dimensão 2) no plano $\mathbb{R}^2$, com conjunto de geradores os vetores $\{v_1, v_2\}$, consiste do conjunto das combinações (lineares) inteiras desses dois vetores, ou seja

$$L = \{v \in \mathbb{R}^2 | v = m_1 v_1 + m_2 v_2,\ m_i \in \mathbb{Z}\}.$$

Definição 3.2. Dados, um reticulado L no plano com conjunto de geradores $\{v_1, v_2\}$, o conjunto dos pontos $a_1 v_1 + a_2 v_2 \in \mathbb{R}^2$ para pos quais $0 \leq a_1 < 1$ é chamado o domínio fundamental do reticulado L associado ao conjunto de geradores $\{v_1, v_2\}$.

Vamos abrir um parênteses para conectar a noção de reticulado e a noção algébrica de grupos. O leitor que não está familiarizado com

a noção de grupos pode, simplesmente utilizar a definição acima, ou ler um pouco do apêndice sobre grupos abelianos. O fato é que se $v, w \in L$, então $v = m_1v_1 + m_2v_2$ e $w = n_1v_1 + n_2v_2$ e, portanto, $v+w = (m_1+n_1)v_1+(m_2+n_2)v_2 \in L$ e $-v = (-m_1)v_1+(-m_2)v_2 \in L$ e estas são as condições para que L seja um subgrupo aditivo de $\mathbb{R}^2$.

Entretanto um reticulado no plano $L \subset \mathbb{R}^2$ não é qualquer tipo de subgrupo. Como conjunto de $\mathbb{R}^2$ um reticulado é, sempre, um subconjunto discreto!!!

Definição 3.3. *Um subconjunto de $X \subset \mathbb{R}^2$ é discreto se todos os seus pontos são isolados, isto é, se dado $p \in X$, existir $\delta > 0$ tal que o único ponto da interseção do disco aberto $D(p, \delta) = \{q \in \mathbb{R}^2 | \ ||q - p|| < \delta\}$ com X for o própio p. Ou seja,*

$$D(p, \delta) \cap X = \{p\}$$

Observação 3.4. Lembramos que os únicos subgrupos discretos da reta real são isomorfos a $\mathbb{Z}$ (e todos são reticulados da reta!!!). De fato, seja $G \subset \mathbb{R}$ um subgrupo aditivo discreto e seja m o menor elemento positivo de G (tal elemento existe pois G é discreto), então $G = m\mathbb{Z}$. (Verifique os detalhes!)

Subgrupos de $\mathbb{R}$ não discretos são bem mais complicados, por exemplo, $\mathbb{Q} \subset \mathbb{R}$ é um subgrupo aditivo que é denso!!!

Proposição 3.5. *Seja $G \subset \mathbb{R}^2$ um subgrupo aditivo. Então são equivalentes:*

1. *G é um reticulado;*
2. *$G \subset \mathbb{R}^2$ é discreto.*

Demonstração: Seja $G \subset \mathbb{R}^2$ um reticulado com conjunto de geradores $\{v, w\}$. Vamos mostrar que G é um conjunto discreto. Vamos mostrar que 0 é um ponto isolado e o resultado segue, por translação. Claramente não existe ponto reticulado no conjunto $int(D) = \{u \in \mathbb{R}^2 | u = \alpha v + \beta w\}$ com $0 < \alpha < 1$ e $0 < \beta < 1$. Assim, tome $\delta = \frac{1}{2}\min\{||v||, ||w||, ||v + w||\}$. Claro que $G \cap D(0, \delta) = 0$.

Reciprocamente, seja $G \subset \mathbb{R}^2$ um subgrupo aditivo discreto. Para mostrar que G é um reticulado devemos encontrar geradores. Seja $v \in G$ o vetor não nulo de menor norma. Seja $w \in G$ o ponto mais próximo da reta $\ell =< v >= \{\lambda v | \lambda \in \mathbb{R}\}$ (não contido na reta). Afirmamos que $G =< v, w >= \{av + bw | a, b \in \mathbb{Z}\}$. Com efeito, seja $u \in G$, e considere os pontos $u - mw \in G$ com $m \in \mathbb{Z}$. O ponto mais próximo da reta ℓ deve pertencer a mesma, caso contrário encontraríamos um ponto mais próximo que w. Assim $u - mw = \lambda v$ e $\lambda = n \in \mathbb{Z}$, logo $u = mv + nw$. □

A partir da proposição acima vemos que é possível exibir um reticulado intrinsecamente, isto é, sem explicitar um conjunto de geradores. Por um lado é mais fácil tratar um reticulado quando conhecemos um conjunto de geradores, por outro lado, muitas vezes é mais fácil provar que um dado conjunto é um reticulado observando que o mesmo é um subgrupo aditivo e discreto do plano $\mathbb{R}^2$.

Exemplo 3.1. Considere o reticulado padrão do plano, ou seja, $L = \mathbb{Z}^2 \subset \mathbb{R}^2$. Temos vários possíveis conjuntos de geradores para tal reticulado. Por exemplo $B_1 = \{(1,0),(0,1)\}$ é um conjunto de geradores para L e $B_2 = \{(2,1),(1,1)\}$ também é um conjunto de geradores para L (faça um esboço dos reticulados associados a estes conjuntos de geradores e verifique que ambos coincidem com $\mathbb{Z}^2$). De fato,

$$(2,1) = 2.(1,0) + 1.(0,1),\ (1,1) = 1.(1,0) + 1.(0,1)$$

logo o reticulado associado a B_2 está contido no reticulado associado a B_1 (combinações inteiras dos vetores de B_2 são combinações inteiras dos vetores de B_1 pois os próprios vetores de B_2 o são) e, reciprocamente

$$(1,0) = 1.(2,1) - 1.(1,1)\ (0,1) = -1.(2,1) + 2.(1,1).$$

Ou seja, os reticulados associados são o mesmo e, claramente, tal reticulado é $\mathbb{Z}^2 \subset \mathbb{R}^2$.

Observação 3.6. A noção de domínio fundamental depende do conjunto de geradores, como mostramos no exemplo anterior. Por outro lado, a área de um domínio fundamental independe do conjunto de geradores. Verifique com exemplos explícitos.

Proposição 3.7. *Sejam $L \subset \mathbb{R}^2$ um reticulado, D e E domínios fundamentais associados, respectivamente, aos conjuntos de geradores $\{v_1, v_2\}$ e $\{u_1, u_2\}$. Então as áreas dos domínios fundamentais são iguais,*

$$A(D) = A(E).$$

Demonstração: Sabemos que existem inteiros $a_1, a_2, b_1, b_2 \in \mathbb{Z}$ tais que

$$u_1 = a_1 v_1 + a_2 v_2 \; u_2 = b_1 v_1 + b_2 v_2$$

pois u_1, u_2 pertencem ao reticulado, logo, são combinação inteira de v_1, v_2 e reciprocamente. Assim, a matriz de mudança de base $M = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$ é inversível e sua inversa N é também uma matriz de coeficientes inteiros(dados pelas coordenadas de v_1, v_2 escritos como combinação inteira de u_1, u_2). Como $M.N = I_2$, $\det(M).\det(N) = 1$ e como são ambos inteiros, $\det(M) = \pm 1$.

A área do domínio fundamental E é

$$A(E) = |\det M| \cdot A(D) = A(D).$$

□

3.2 O Toro Plano

Como vimos na seção anterior, proposição 3.7, a área de um domínio fundamental de um reticulado $L \subset \mathbb{R}^2$ é um invariante que depende somente do reticulado e não da escolha de geradores. Nessa seção vamos conectar a álgebra com a geometria e tratar mais a fundo a noção de área. A noção de área esbarra em dois problemas sérios: mensurabilidade e métrica. O primeiro é já um problema no plano, onde existem conjuntos que não possuem uma área bem definida, por exemplo o conjunto abaixo do gráfico de uma função não integrável. O segundo é um problema que fica mais evidente quando se pretende definir área em uma superfície abstrata, a métrica é o ingrediente necessário para isso. Nessa seção vamos lidar com uma superfície

interessante cujo modelo mais conhecido, mergulhado em $\mathbb{R}^3$, possui uma determinada métrica(proveniente da métrica usual do $\mathbb{R}^3$) mas no nosso caso vamos "exportar"a métrica do plano para "medir"áreas nessa superfície. A superfície é o TORO e munida da métrica "do plano"é chamada *O Toro Plano.*

Definição 3.8. *Seja $L \subset \mathbb{R}^2$ um reticulado plano. O grupo (de Lie) quociente $\mathbb{R}^2/L$ é a superfície que chamaremos o toro plano.*

É claro que, conjuntisticamente, o quociente $\mathbb{R}^2/L$ pode ser identificado com um domínio fundamental e, a partir desta identificação é que exportamos a métrica do plano ao toro. Não é esse o momento para entrar em detalhes técnicos (que são muitos) sobre a estrutura algebro-topológico-geométrica do toro. Vamos mostrar que, como grupo (de Lie), o toro é isomorfo ao produto de dois círculos. Consideramos o círculo $\mathbb{S}^1 \subset \mathbb{R}^2$ como $\mathbb{S}^1 = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 = 1\}$ ou, equivalentemente, $\mathbb{S}^1 = \{z \in \mathbb{C} | \ |z| = 1\}$. Lembramos que $\mathbb{S}^1$ possui uma estrutura de grupo. Usaremos parametrizações clássicas, com a exponencial complexa ou com cossenos e senos no caso real.

Proposição 3.9. *Seja $L \subset \mathbb{R}^2$ um reticulado no plano e $D \subset \mathbb{R}^2$. Então o quociente $\mathbb{R}^2/L$ pode ser identificado com o toro $T = \mathbb{S}^1 \times \mathbb{S}^1$. Mais precisamente existe um isomorfismo de grupos (de Lie) entre eles.*

Demonstração: Sejam $\{v_1, v_2\}$ um sistema de geradores para L. Considere o seguinte homomorfismo de grupos:

$$\begin{array}{rccc} \phi: & \mathbb{R}^2 & \to & T = \mathbb{S}^1 \times \mathbb{S}^1 \\ & (a_1 v_1 + a_2 v_2) & \mapsto & (e^{2\pi i a_1}, e^{2\pi i a_2}) \end{array}$$

em que

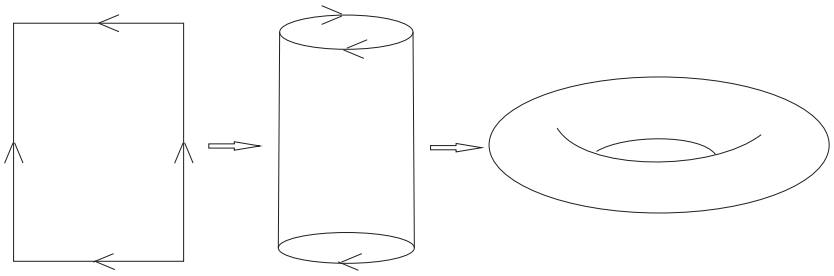
$$(e^{2\pi i a_1}, e^{2\pi i a_2}) = (((cos(2\pi a_1), sen(2\pi a_1)); (cos(2\pi a_2), sen(2\pi a_2)))$$

claramente esse é um hoomomorfismo sobrejetivo (pois utilizamos uma parametrfização do toro) e seu núcleo é L. Logo, pelo teorema do isomorfismo concluímos que

$$\mathbb{R}^2/L \cong T.$$

□

Geometricamente o isomorfismo explicitado no teorema anterior corresponde à famosa identificação



Definição 3.10. *Seja $X \subset T$ uma região no toro $T = \mathbb{R}^2/L$ associado ao reticulado $L \subset \mathbb{R}^2$ e seja D um domínio fundamental para L. Definimos a área desta região por*

$$A(X) = A(\phi|_D{}^{-1}(X))$$

desde que exista $A(\phi|_D{}^{-1}(X))$, que é a área de uma região no plano (desde que exista a área dessa região no plano).

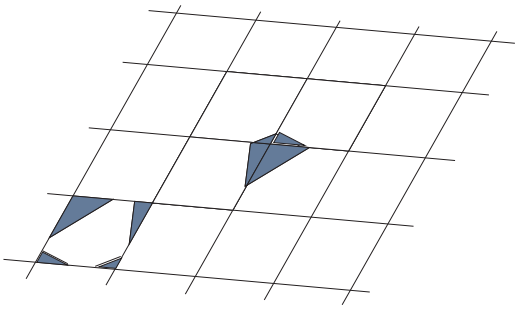
Observação 3.11. *Observamos que a definição acima independe de escolha de domínio fundamental e que, em particular, a área total do toro é igual à área de um domínio fundamental que, pela proposição 3.7, independe do domínio fundamental.*

Proposição 3.12. *Se $Y \subseteq \mathbb{R}^2$ é limitada e existe $A(Y)$, e se $A(\phi(Y)) \neq A(Y)$ então $\phi|_Y$ não é injetiva.*

Demonstração: Supondo que $\phi|_Y$ seja injetiva, $Y = \bigcup Y_i$, onde $Y_i = Y \cap (D + w_i)$ são disjuntos. Fazendo $Z_i = Y_i - w_i \subset D$ são também disjuntos pela injetividade de $\phi|_Y$ logo:

$$A(\phi(Y)) = A(\phi(\bigcup Y_i)) = A(\bigcup Z_i) = \sum A(Z_i) = \sum A(Y_i) = A(Y).$$

□



3.3 Reticulados Inteiros no Plano

Trataremos, agora, um tipo especial de reticulado. Chamaremos de reticulado inteiro no plano um (sub)reticulado $L \subset \mathbb{Z}^2$ do reticulado padrão. É interessante notar que se $L \subset \mathbb{Z}^2$, então L é necessariamente discreto, portanto, basta que seja um subgrupo aditivo, para que seja um reticulado no plano.

O próximo resultado nos será útil para cálculos efetivos e relaciona a área de um domínio fundamental de um reticulado inteiro no plano $A(D)$, o número de elementos do grupo quociente $|\mathbb{Z}^2/L|$ e o número de pontos inteiros em D. Todas as quantidades citadas coincidem!!!

Proposição 3.13. *Dados um reticulado $L \subset \mathbb{Z}^2$ e D um domínio fundamental. A área de D será*

$$A(D) = |\mathbb{Z}^2/L|,$$

que corresponde ao número de pontos inteiros em um domínio fundamental.

Demonstração: É claro que $|\mathbb{Z}^2/L|$ é igual ao número de pontos inteiros em um domínio fundamental. Com efeito, se L é um reticulado e D é um domínio fundamental, então todo vetor do plano pertence a exatamente um conjunto $D + w$, em que w é um ponto do reticulado.

Se mostrarmos que a área de D também é igual ao número de pontos inteiros de um domínio fundamental, então o resultado segue. A demonstração de que a área do domínio fundamental é igual ao número de pontos inteiros do domínio fundamental segue imediatamente do teorema de Pick, teorema 1.9, (verifique!). □

3.4 Teorema de Minkowski

Nessa seção apresentamos o principal resultado técnico deste minicurso. A idéia intuitiva associada a este resultado é relativamente simples e explicaremos agora. Para isso precisaremos de dois conceitos geométricos:

Definição 3.14. *Um subconjunto do plano $X \subset \mathbb{R}^2$ é dito ser convexo se dados dois pontos $x, y \in X$ o segmento de reta unindo esses pontos $[x, y] = \{ax + (1-a)y | a \in [0, 1]\}$ está completamente contida em X, isto é, $[x, y] \subset X$.*

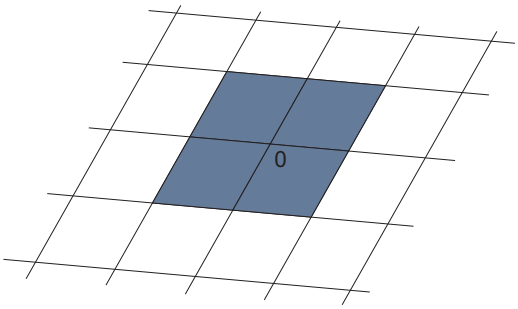
Observação 3.15. *Um polígono no plano é convexo se, e somente se, cada um de seus ângulos internos for menor que π. São ainda exemplos de subconjuntos convexos do plano o disco, a região interior de uma elipse (e também o seu fecho),...*

Definição 3.16. *Um subconjunto do plano $X \subset \mathbb{R}^2$ será dito ser simétrico (com relação à origem) se para cada $x \in X$ tivermos $-x \in X$.*

Observação 3.17. *Paralelogramos, polígonos regulares e discos (centrados na origem) são exemplos triviais de subconjuntos limitados simétricos e convexos em $\mathbb{R}^2$.*

Dado um reticulado no plano $L \subset \mathbb{R}^2$ e D um domínio fundamental (associado a um conjunto de geradores $\{v_1, v_2\}$). O mais simples conjunto convexo e simétrico do plano, associado a L, com área máxima e sem conter pontos não nulos do reticulado são:

paralelogramos "semelhantes ao paralelogramo gerado pelos vetores v_1 e v_2 "e centrados na origem.



É claro que se quisermos nos esquivar dos pontos não nulos do reticulado tais paralelogramos devem ter área menor que 4 vezes a área do domínio fundamental. O teorema de Minkowski formaliza essa idéia mas não somente para os (intuitivos) paralelogramos semelhantes ao domínio fundamental e sim para qualquer conjunto limitado simétrico e convexo. Precisamente, temos o seguinte:

Teorema 3.18. *Sejam L um reticulado em $\mathbb{R}^2$, D um domínio fundamental para L e $X \subset R^2$, um conjunto limitado, simétrico e convexo tal que $A(X) > 4A(D)$. Então X contém um ponto não nulo de L.*

Demonstração: Duplicando L, obtém-se um reticulado $2L$ com domínio fundamental $2D$ cuja área é $4A(D)$. Considerando o toro relativo a tal reticulado:

$$T = \mathbb{R}^2/2L$$

Cuja área é $A(T) = A(2D) = 4A(D)$.

Logo $\phi : \mathbb{R}^2 \to T$, o homomorfismo estrutural não preserva a área de X. Pois:

$$A(\phi(X)) \leq A(T) = 4A(D) < A(X).$$

Então, pela proposição 3.12, $\phi|_X$ não é injetiva. Assim, existem $x_1 \neq x_2$, $x_1, x_2 \in X$, tais que:

$$\phi(x_1) = \phi(x_2)$$

ou equivalentemente, $x_1 - x_2 \in 2L = Ker(\phi)$.

Por X ser simétrico tem-se que $x_2 \in X$ o que implica que $-x_2 \in X$;e como X é convexo $\frac{1}{2}(x_1)+\frac{1}{2}(-x_2) \in X$, ou seja, $\frac{1}{2}(x_1-x_2) \in X$.

Portanto:

$$\frac{1}{2}(x_1 - x_2) \in L \cap X.$$

Este é um ponto não nulo do reticulado L e que pertence a X. □

3.5 Problemas

1. Considere os reticulados, dados por um conjunto de geradores $\{u, v\}$:
   (a) $u = (1, 2)$ e $v = (1, 1)$
   (b) $u = (2, 2)$ e $v = (1, 3)$
   (c) $u = (1, 2)$ e $v = (-1, 1)$
   (d) $u = (1, \pi)$ e $v = (-1, \pi)$
2. Prove que o disco $D = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 \leq r^2\}$ é convexo.
3. Prove que o quadrado $Q = \{(x, y) \in \mathbb{R}^2 |\ |X| \leq c\,|y| \leq c\}$ é convexo.
4. Generalizar todas as definições e resultados desse capítulo para reticulados em $\mathbb{R}^3$. Observamos que tudo pode ser ainda generalizado para o $\mathbb{R}^n$ como em [1].

Capítulo 4

Soma de Dois Quadrados

4.1 Introdução

O teorema dos Dois Quadrados afirma que os únicos primos p que podem ser escritos como soma de dois quadrados de inteiros, $p = x^2 + y^2$, com $x, y \in \mathbb{Z}$ são $p = 2$ e os primos da forma $p = 4k + 1$. Esse teorema, juntamente com outros dois de mesma natureza foram descobertos por Fermat. Em 1640 Fermat enviou uma carta a Mersene com o enunciado do teorema, em 1659 ele enviou uma carta a Pierre de Carcavi com um esboço da prova. Em 1754 Euler fornece uma demonstração completa do teorema. Euler esteve 40 anos de sua vida estudando esses problemas de Fermat sobre primos da forma $x^2 + ny^2$.
São problemas interessantes também o de soma de três quadrados, soma de quatro quadrados e o geral que é conhecido como problema de Waring. Identificar os inteiros que podem ser escritos como uma soma de n quadrados de inteiros.

4.2 Pontos Inteiros VS Pontos Racionais em Círculos

Na seção intitulada Aritmética em Cônicas enunciamos um importante Teorema de Hasse-Minkowsky, Teorema 2.11, que dava condições necessárias e suficientes para que uma cônica com coeficientes racionais possua ponto(s) racional(is). Um exemplo que foi analisado particularmente foi a cônica:

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

que mostramos não possuir ponto racional. O teorema a seguir mostra que, para esse tipo de equações, é suficiente verificar a não existência de pontos inteiros (que é muitíssimo mais fácil!!!) para concluir a não existência de pontos racionais.

Teorema 4.1. *Seja $n \in \mathbb{Z}$ um inteiro positivo, então n é soma de dois quadrados de racionais se, e somente se, n for soma de dois quadrados de inteiros.*

Demonstração: Suponhamos que n seja soma de dois quadrados de racionais $p_1^2 + p_2^2 = n$ com $p_1 \notin \mathbb{Z}$ ou $p_2 \notin \mathbb{Z}$. Seja $P = (p_1, p_2)$ o ponto do círculo $x^2 + y^2 = n$. Seja $M = (m_1, m_2) \in \mathbb{Z}^2$ o ponto inteiro tal que $|m_i - p_i| \leq \frac{1}{2}$ $i = 1, 2$. A reta $\ell = \overline{MP}$ não pode ser tangente ao cículo $x^2 + y^2 = n$. Com efeito, se ℓ fosse tangente ao círculo, então o triângulo OPM seria retângulo em P (ponto de tangência). Assim $\overline{OM}^2 = \overline{OP}^2 + \overline{PM}^2$, e isso é um absurdo pois $\overline{OM}^2 \in \mathbb{Z}$, $\overline{OP}^2 = p_1^2 + p_2^2 = n \in \mathbb{Z}$ (por hipótese) e $0 \neq \overline{PM}^2 = |m_1 - p_1|^2 + |m_2 - p_2|^2 \leq \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$.

Logo, a reta $\overline{PM}$ é secante ao círculo, ambos(a reta e o círculo) possuem coeficientes racionais e se intersectam em P ponto racional. Pelo método de Fermat, o outro ponto de interseção, $Q = (q_1, q_2)$ é também racional.

Seja d o mmc das frações irredutíveis p_1, p_2 que definem P. Defina $c = d|PM|^2 < d$,

$$c = d(|m_1 - p_1|^2 + |m_2 - p_2|^2) = d[m_1^2 + m_2^2 + n - 2(p_1 m_1 + p_2 m_2)] \in \mathbb{Z} \tag{4.1}$$

Vamos mostrar que c elimina os denominadores de q_1 e q_2. Isso conclui a prova pois, a partir de P obtemos Q e reduzimos os denominadores, se procedermos assim, em algum momento encontraremos um ponto inteiro.

Ora, $Q = P + t(M - P) = (p_1 + t(m_1 - p_1), p_2 + t(m_2 - p_2))$ com $t \in \mathbb{Q}^*$. Defina $v = M - P = (m_1 - p_1, m_2 - p_2)$. Como Q pertence ao círculo de equação $x^2 + y^2 = n$, temos que $Q \cdot Q = n$ (produto escalar é indicado por $\cdot$). Logo:

$$n = (P + tv) \cdot (P + tv) = P \cdot P + 2t(P \cdot v) + t^2(v \cdot v)$$

como $P \cdot P = n$ (P é um ponto do círculo), então $2t(P \cdot v) + t^2(v \cdot v) = 0$ e como $v \cdot v = ||PM||^2 = \frac{c}{d}$, temos

$$t = -2\frac{P.v}{v.v} = -2\frac{p_1m_1 + p_2m_2 - n}{\frac{c}{d}} = \frac{d(2n - 2(p_1m_1 + p_2m_2))}{c}. \tag{4.2}$$

Vamos finalmente mostrar que c elimina os denominadores de q_1 e q_2. Devemos mostrar que

$$cq_i = c(p_i + t(m_i - p_i)) = cp_i + (ct)(m_i - p_i) \tag{4.3}$$

é inteiro.

Das equações 4.2 e 4.2, temos que

$$ct = d(2n - 2(p_1m_1 + p_2m_2)) = d(2n + \frac{c}{d} - n - m_1^2 - m_2^2) = c + d(n - m_1^2 - m_2^2).$$

Substituindo em 4.3 temos

$$cq_i = cp_i + [c + d(n - m_1^2 - m_2^2)](m_i - p_i) = cm_i + d(n - m_1^2 - m_2^2)(m_i - p_i).$$

Claro que esses números são inteiros pois c, m_i e n são inteiros e d elimina os denominadores de p_1 e p_2.

O resultado segue, pois, escolhendo P (ponto racional do círculo) de forma que o mmc entre os denominadores de p_1 e p_2 fosse mínimo, teríamos encontrado Q (ponto racional do círculo) cujo mmc dos denominadores seria menor. □

Observação 4.2. Novamente a técnica de demonstração desse teorema é a "descida infinita de Fermat". Essa técnica é aplicada em vários problemas aritméticos.

4.3 Inteiros de Gauss

O conjunto $\mathbb{Z}[i] = \{a + bi | a, b \in \mathbb{Z}\} \subset \mathbb{C}$ em que $i^2 = -1$ é chamado conjunto dos inteiros gaussianos. Esse conjunto foi estudado por Gauss (daí o nome) e é muito semelhante ao conjunto $\mathbb{Z}$ dos números inteiros, tanto do ponto de vista algébrico (propriedades da adição e multiplicação, ...) quanto do ponto de vista aritmético (algoritmo de divisão, mdc, fatoração...).

Para nós, entretanto, será interessante um único aspecto de tal conjunto, que é derivado das noções de conjugação e norma que existem no corpo dos números complexos.

Definição 4.3. *A conjugação em $\mathbb{Z}[i]$ é definida da seguinte forma:*

$$\overline{a + bi} = a - bi.$$

Proposição 4.4. *A conjugação em $\mathbb{Z}[i]$ satisfaz as seguintes propriedades:*

(i) $\overline{z + w} = \overline{z} + \overline{w}$;

(ii) $\overline{z.w} = \overline{z}.\overline{w}$;

(iii) $z\overline{z} \geq 0$ *e* $z\overline{z} \in \mathbb{Z}$.

Demonstração: Verifique!!! □

A partir da idéia de conjugação e pelo item (iii) da proposição anterior, podemos definir a chamada norma algébrica em $\mathbb{Z}[i]$.

Definição 4.5. *Definimos a norma algébrica de um elemento $z = a + bi \in \mathbb{Z}[i]$ por*

$$N(z) = z\overline{z} = a^2 + b^2 \in \mathbb{Z}.$$

Proposição 4.6. *Sejam $z, w \in \mathbb{Z}[i]$, então*

$$N(zw) = N(z)N(w).$$

Ou seja, se $m, n \in \mathbb{Z}$ são dois inteiros positivos que são norma de elementos de $\mathbb{Z}[i]$, isto é, se $n = N(z)$ e $m = N(w)$, então mn também é norma de algum elemento de $\mathbb{Z}[i]$, de fato, $mn = N(zw)$.

4.4 Soma de Dois Quadrados

Consideramos agora, para n um natural, a equação

$$x^2 + y^2 = n$$

Quando a equação tem solução natural dizemos que n é soma de dois quadrados. Nosso objetivo é caracterizar os naturais que são soma de dois quadrados.

Lema 4.7. *Seja $p \equiv 1 \pmod 4$ um número primo. Então existe $u \in \mathbb{Z}/p\mathbb{Z}$ tal que $u^2 \equiv -1 \in \mathbb{Z}/p\mathbb{Z}$.*

Demonstração: Pelo pequeno Teorema de Fermat sabemos que todo elemento de $\mathbb{Z}/p\mathbb{Z}$ não nulo satisfaz

$$a^{p-1} = \overline{1}.$$

Ora, $p = 4k + 1$, considere a equação:

$$x^{2k} = -\overline{1} \in \mathbb{Z}/p\mathbb{Z}.$$

Mostraremos que essa equação possui solução em $\mathbb{Z}/p\mathbb{Z}$ e assim o resultado segue, pois, se w é solução de tal equação, então $u = w^k$ satisfaz nosso enunciado, de fato, $u^2 = w^{2k} = -\overline{1} \in \mathbb{Z}/p\mathbb{Z}$.

Ora, sendo que $a^{4k} = \overline{1}$ para todo $a \in \mathbb{Z}/p\mathbb{Z}$, $a \neq \overline{0}$, temos que a equação

$$x^{4k} = \overline{1} \in \mathbb{Z}/p\mathbb{Z}$$

possui $4k = p - 1$ soluções. Mas $x^{4k} = \overline{1} \Rightarrow (x^{2k} - 1)(x^{2k} + 1) = 0$ e assim temos duas possibilidades: $x^{2k} - 1 = 0$ ou $x^{2k} + 1 = 0$. Como

$\mathbb{Z}/p\mathbb{Z}$ é um corpo, então não pode ocorrer da equação $x^{2k} - 1 = 0$, de grau $2k$, possuir $4k$ soluções. Logo, a equação $x^{2k} + 1 = 0$ possui alguma solução e o resultado segue. □

O resultado a seguir possui muitas demonstrações diferentes na literatura. A demonstração que daremos segue de [1] e é bem geométrica, utilizando o teorema de Minkowski.

Teorema 4.8 (Fermat-Euler). *Um primo $p > 0$ é soma de dois quadrados de inteiros se, e somente se, $p = 2$ ou p é da forma $4k+1$, $k \in \mathbb{Z}$.*

Demonstração: É claro que $2 = 1^2 + 1^2$ é soma de dois quadrados de inteiros. Portanto só nos resta provar o resultado para $p \neq 2$.

Se $p \neq 2$ e $p \not\equiv 1 \pmod 4$, então $p \equiv 3 \pmod 4$. Suponha que existam dois inteiros a e b tais que $a^2 + b^2 = p$. Fazendo congruência módulo 4 temos: $a^2 + b^2 \equiv 3 \pmod 4$ que é um absurdo (verifique!).

Reciprocamente, se $p \equiv 1 \pmod 4$, então mostraremos que p é soma de dois quadrados de inteiros. Pelo lema 4.7, existe $u \in \mathbb{Z}/p\mathbb{Z}$ tal que $u^2 \equiv -1 \in \mathbb{Z}/p\mathbb{Z}$. Consideremos, agora, o reticulado $L = \{(a, b) \in \mathbb{Z}^2 | b = ua \in \mathbb{Z}/p\mathbb{Z}\}$. Analisando a aplicação:

$$\begin{array}{ccc} \mathbb{Z}^2 & \twoheadrightarrow & \mathbb{Z}/p\mathbb{Z} \\ (x, y) & \mapsto & y - \bar{u}x \end{array}$$

nota-se que $ker(y) = L$ e, pelo teorema do Isomorfismo, $\mathbb{Z}^2/L \simeq \mathbb{Z}/p\mathbb{Z}$. Logo, pela proposição 3.13, a área de um domínio fundamental D, do reticulado L é $A(D) = p$.

Considere agora $X \subset \mathbb{R}^2$, o disco limitado, simétrico e convexo

$$X = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 \leq \frac{3p}{2}\}.$$

Como a área de X é $\pi r^2 = \pi \frac{3p}{2} > 4p$, então, pelo teorema de Minkowski, teorema 3.18, existe um ponto $0 \neq (a, b) \in L \cap X$. Ou seja

$$0 \neq a^2 + b^2 \leq r^2 = \frac{3p}{2} < 2p$$

analisando $a^2 + b^2$ módulo p e lembrando que $u^2 \equiv -1 \pmod{p}$, obtemos:

$$a^2 + b^2 \equiv a^2 + u^2a^2 \equiv a^2 - a^2 \equiv 0 \pmod{p}$$

Logo,

$$a^2 + b^2 = p$$

Pois p é o único múltiplo não nulo de p e menor que $2p$.

Logo, se $p \equiv 1 \pmod{4}$, então p é soma de dois quadrados de inteiros. □

Teorema 4.9. *Seja $n > 0$ um inteiro. Então n é soma de dois quadrados de inteiros se, e somente se, n* **não** *possui, em sua fatoração em primos, uma potência de expoente ímpar para um primo $p \equiv 3 \pmod{4}$. Ou seja, na fatoração de n podem ocorrer os primos 2 e $p \equiv 1 \pmod{4}$ com expoente arbitrário, mas os primos da forma $p \equiv 3 \pmod{4}$ devem ocorrer com expoente par.*

Demonstração: Suponhamos que

$$n = p_1^{e_1} p_2^{e_2} ... p_k^{e_k}$$

é a fatoração em primos (distintos) de n e suponhamos que todos os primos $p \equiv 3 \pmod{4}$ que ocorrem na fatoração de n possuem expoente par. Para cada primo $p = 2$ ou $p \equiv 1 \pmod{4}$, p é soma de dois quadrados, pelo teorema 4.8 e, portanto, qualquer potência de tais primos é também soma de dois quadrados, pela proposição 4.6. Para os primos que possuem expoente par, então a própria potência é um quadrado e portanto soma de quadrados ($x^2 = 0^2 + x^2$). O resultado segue do fato que o produto de inteiros que são soma de dois quadrados é também soma de dois quadrados novamente, pela proposição 4.6.

Reciprocamente, suponhamos que $n = a^2 + b^2$ seja soma de dois quadrados. Suponhamos que n possua algum fator primo, $p \equiv 3 \pmod{4}$, cujo expoente em sua fatoração (em primos distintos) seja

ímpar. Seja $d = mdc(a, b)$. Como $d^2|a^2 + b^2 = n$ e $p|n$ (com multiplicidade ímpar), então $p|\frac{n}{d^2} = m$. Sejam u e v tais que $a = du$ e $b = dv$, então $u^2 + v^2 = m$, $mdc(u, v) = 1$ e $p|m$. Assim, existe um dentre os números u e v que é coprimo com p (digamos u) e, portanto, existe w tal que $uw \equiv 1 \pmod{p}$. Como $u^2 + v^2 \equiv 0 \pmod{p}$ temos que $(wv)^2 \equiv -1 \pmod{p}$. Ora, $p \equiv 3 \pmod{4}$ logo $p - 1 = 2q$ em que q é ímpar. Considere a expressão $(wv)^2 \equiv -1 \pmod{p}$ tomando potências com expoente q, obtemos $(wv)^{p-1} \equiv -1 \pmod{p}$ (lembramos que q é ímpar!). Isso é um absurdo, pelo pequeno teorema de Fermat. □

4.5 Problemas

1. Seja $p \equiv 1 \pmod{4}$ um primo positivo. Mostre que o círculo $x^2 + y^2 = p$ possui um único ponto inteiro e positivo.

2. Dê exemplos de números inteiros e positivos n para os quais o círculo $x^2 + y^2 = n$ possui mais de um ponto inteiro e positivo.

3. Mostre que para um primo $p > 0$ são equivalentes:

   (i) Existe $u \in \mathbb{Z}$ tal que $u^2 \equiv -2 \pmod{p}$

   (ii) ...

4. Mostre que para a classe de primos descrita no problema anterior são precisamente os primos que podem ser escritos da forma $p = x^2 + 2y^2$, com $x, y \in \mathbb{Z}$. Dica: Imite a demonstração do teorema 4.8.

5. Mostre que um inteiro positivo n é soma de três quadrados de inteiros se, e somente se, n é soma de três quadrados de racionais.

6. Mostre que todo número inteiro positivo n pode ser escrito como soma de quatro quadrados de inteiros. Sugestões:

   (a) Use a norma dos quartenions para se reduzir ao caso de $n = p$ primo.

(b) Mostre que a congruência $u^2 + v^2 + 1 \equiv 0 \pmod{p}$ possui solução qualquer que seja $p > 0$ primo.

(c) Escolha $u, v \in \mathbb{Z}$ como no item anterior e defina o seguinte reticulado: $L \subset \mathbb{Z}^4$ definido por $L = \{(a, b, c, d) \in \mathbb{Z}^4 | c = au + bv, d = ub - va\}$

(d) Utilize o teorema do isomorfismo para concluir que o volume de um domínio fundamental de L é p^2

(e) Utilize o teorema de Minkowski em $\mathbb{R}^4$. Lembre-se que o volume da esfera em $\mathbb{R}^4$ é $V = \frac{\pi^2 r^2}{2}$

Capítulo 5

Equações de Pell-Fermat

5.1 Introdução

Seja $d \in \mathbb{Z}$ um inteiro positivo que não é quadrado. Chamamos equação de Pell-Fermat as equações do tipo

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

das quais se procuram soluções inteiras. Tais equações representam hipérboles que mostraremos possuir uma infinidade de pontos inteiros.

As equações de Pell-Fermat são estudadas ha milênios na Índia e na Grécia. Eles estavam particularmente interessados no caso $d = 2$ uma vez que suas soluções forneciam boas aproximações racionais de $\sqrt{2} \simeq \frac{x}{y}$. Baudhayana (800 AC) encontrou os pares $(17, 12)$ e $(577, 408)$ que forneciam muito boas aproximações para $\sqrt{2} \simeq \frac{577}{408}$. Arquimedes (300 AC) usou a equação no caso $d = 3$ e obteve a aproximação $\sqrt{3} \simeq \frac{1351}{780}$. Brahmagupta e Bhaskara também estudaram essas equações utilizando um método denominado *Chakravala*.

O nome de Pell, nestas equações, ocorre devido a um erro de Euler atribuindo ao matemático inglês John Pell (1610-1685) o estudo da mesma. Aparentemente foi Lord Brouncker (1620-1684) o primeiro matemático europeu moderno a estudar as equações de Pell-Fermat. Fermat, em 1657 propôs o problema geral de resolver tais equações em uma carta a Frenicle. Euler, em 1770 estuda estas equações em seu livro de Álgebra utilizando frações contínuas. Lagrange, em 1773 demonstra o teorema 5.6. Nosso enfoque será mais geométrico, em particular, daremos uma demonstração via Teorema de Minkowski da proposição chave para mostrar que as equações de Pell-Fermat sempre possuem uma infinidade de soluções, Proposição 5.7.

5.2 Inteiros Quadráticos de Pell-Fermat

Uma álgebra sobre os inteiros (que seja domínio de integridade) é chamada um Domínio de Inteiros Quadráticos se $A \simeq \mathbb{Z}[X]/(f)$ em que f é um polinômio irredutível de grau 2. Nós estaremos interessados num tipo um pouco mais simples e suporemos que $f = X^2 - d$ em que d é um inteiro positivo não quadrado. Pelo teorema do isomorfismo temos que tais conjuntos são da forma

$$A \simeq \mathbb{Z}[\sqrt{d}] = \{a + b\sqrt{d} | a, b \in \mathbb{Z}\} \subset \mathbb{C}$$

tomaremos então essa descrição para os nossos domínios quadráticos de Pell-Fermat.

Definição 5.1. *Um Domínio de Inteiros Quadráticos de Pell-Fermat é um conjunto da forma*

$$\mathbb{Z}[\sqrt{d}] = \{a + b\sqrt{d} | a, b \in \mathbb{Z}\} \subset \mathbb{R}$$

em que $d \in \mathbb{Z}$ é um inteiro positivo não quadrado.

Dizer que esse conjunto é um domínio (de integridade) significa dizer que o mesmo está munido de duas operações, adição e multiplicação, satisfazendo as propriedades usuais (como $\mathbb{Z}$). Não exigimos a existência de inversos multiplicativos em geral, somente a propriedade que caracteriza os domínios. Se $xy = 0$, então $x = 0$ ou $y = 0$. Há, para esses domínios, uma noção de conjugação muito similar a conjugação complexa.

Definição 5.2. *Seja $z = a + b\sqrt{d} \in \mathbb{Z}[\sqrt{d}]$, o conjugado de z é $\overline{z} = a - b\sqrt{d}$.*

A conjugação em $\mathbb{Z}[\sqrt{d}]$ satisfaz as mesmas propriedades da conjugação complexa, que são

1. $\overline{z+w} = \overline{z} + \overline{w}$;
2. $\overline{z.w} = \overline{z}.\overline{w}$

Definição 5.3. *A norma algébrica em $\mathbb{Z}[\sqrt{d}]$ é definida por:*

$$N(z) = z\overline{z}$$

logo, se $z = a + b\sqrt{d}$, então $N(z) = a^2 - b^2 d \in \mathbb{Z}$.

A propriedade fundamental da norma algébrica de $\mathbb{Z}[\sqrt{d}]$ é

$$N(zw) = N(z)N(w)$$

Proposição 5.4. *Um elemento $z \in \mathbb{Z}[\sqrt{d}]$ é inversível se, e somente se, $N(z) = \pm 1$.*

Demonstração: Se z é inversível, então existe $w \in \mathbb{Z}[\sqrt{d}]$ tal que $zw = 1$ e, portanto, $N(z)N(w) = 1$ como $N(z)$ e $N(w)$ são números inteiros, temos que $N(z) = \pm 1$ (que são os únicos inteiros inversíveis).

Reciprocamente, se $N(z) = \pm 1$, então $z\overline{z} = \pm 1$ e portanto $z(\mp\overline{z}) = 1$ donde concluímos que z é inversível. □

Observação 5.5. *Se $\in \mathbb{Z}[\sqrt{d}]$ é inversível, então $z^{-1} = \overline{z}$ pois, nesse caso, $N(z) = z.\overline{z} = 1$. Se $z \in \mathbb{Z}[\sqrt{d}]$ é inversível e $N(z) = 1$, então $N(z^{-1}) = 1$, ou seja, se $z = x + y\sqrt{d} = 1$. Então $z^{-1} = \overline{z} = x - y\sqrt{d} = 1$.*

Portanto, a partir desta proposição podemos concluir que são equivalentes quando $d > 0$:

(i) $z = a + b\sqrt{d} > 0$ com $a, b > 0$ é inversível em $\mathbb{Z}[\sqrt{d}]$;

(ii) (a, b) é uma solução em inteiros positivos da equação $x^2 - y^2 d = 1$;

(iii) $z = a + b\sqrt{d} > 0$ com $a, b > 0$ e $N(z) = 1$.

O ponto que nos interessa é que se $z = a + b\sqrt{d} \neq \pm 1$, então, não apenas ele, bem como todas as suas potências dão origem à soluções não triviais da equação $x^2 - dy^2 = 1$, pela proposição 4.6. Nesse caso as potências de z nos fornecem uma infinidade de soluções para a equação $x^2 - dy^2 = 1$ desde que $z \neq \pm 1$.

5.3 Soluções da Equação de Pell-Fermat

Seja $d \in \mathbb{Z}$ um inteiro não quadrado. As equações do tipo

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

são chamadas equações de Pell-Fermat. Vamos, agora, utilizar a Teoria de Minkowski sobre reticulados no plano para mostrar que as hipérboles definidas por uma equação de Pell-Fermat sempre possuem uma infinidade de pontos inteiros. Chamamos soluções triviais $(\pm 1, 0)$. Podemos nos reduzir a procura de soluções positivas, isto é, $x > 0$ e $y > 0$ uma vez que todas as outras soluções inteiras são obtidas a partir destas.

Teorema 5.6. *Seja $d \in \mathbb{Z}$ um inteiro positivo não quadrado, então a equação de Pell-Fermat*

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

possui uma infinidade de soluções inteiras positivas. Além disso, se (x_1, y_1), $x_1 > 0$ e $y_1 > 0$, é a solução tal que $x_1 + y_1\sqrt{d}$ é mínimo, então todas as outras soluções inteiras positivas da equação de Pell-Fermat são (x_n, y_n) tal que $x_n + y_n\sqrt{d} = (x_1 + y_1\sqrt{d})^n$

De fato, a partir da discussão do fim da última seção, percebemos que é suficiente mostrar que existe uma solução $(a, b) \in \mathbb{Z}^2$ tal que $a + b\sqrt{d} > 1$. Com efeito, sendo $z_1 = a + b\sqrt{d} \in \mathbb{Z}[\sqrt{d}]$ definimos $z_k = z_1^k$. Como $z_1 > 1$, então todas as suas potências são distintas e além disso, (a, b) é solução da equação de Pell-Fermat se, e somente se, $N(a + b\sqrt{d}) = 1$. Nesse caso, não apenas z_1 é de norma 1 como também todas as suas potências. E assim obtemos um infinidade de soluções.

Exemplo 5.1. Considere a equação

$$x^2 - 2y^2 = 1.$$

Uma solução não trivial é $(3, 2)$ que corresponde ao inversível $z = 3+2\sqrt{2} \in \mathbb{Z}[\sqrt{2}]$. Suas potências são: $z^2 = 17+12\sqrt{2}$, $z^3 = 99+70\sqrt{2}$, $z^4 = 577 + 408\sqrt{2}$, ...

Esses elementos de $\mathbb{Z}[d]$ dão origem às seguintes soluções da equação de Pell-Fermat:$(\pm 3, \pm 2)$ $(\pm 17, \pm 12)$, $(\pm 99, \pm 70)$, $(\pm 577, \pm 408)$, ... Pelo teorema de Pell Fermat essas são todas as soluções da equação.

A próxima proposição será fundamental para a demonstração do Teorema de Pell-Fermat e apresentamos aqui uma demonstração geométrica original. As demonstrações anteriores usavam teoria de aproximação.

Proposição 5.7. *Seja $d \in \mathbb{Z}$ um inteiro positivo não quadrado. Então existe $m > 0$ tal que a equação*

$$x^2 - dy^2 = m$$

possui uma infinidade de soluções inteiras.

Demonstração: Considere, primeiramente, o reticulado

$$L_1 = \{(a + b\sqrt{d}, a - b\sqrt{d}) \in \mathbb{R}^2 | a, b \in \mathbb{Z}\}.$$

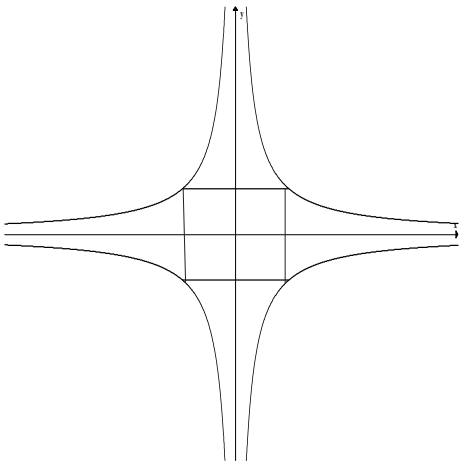
Um conjunto de geradores pala L_1 é $\{(1, 1), (\sqrt{d}, -\sqrt{d})\}$, logo, a área de um domínio fundamental de L_1 é $A = |\det((1, 1), (\sqrt{d}, -\sqrt{d}))| = 2\sqrt{d}$. Observamos que se $(x, y) = (a + b\sqrt{d}, a - b\sqrt{d}) \in L_1$, então $xy = a^2 - db^2 = N(a + b\sqrt{d})$. Nosso objetivo, portanto, será encontrar elementos cujo produto das coordenadas seja limitado.

Considere ainda, para cada $c > 0$ os seguintes conjuntos:

$$H_c = \{(x, y) \in \mathbb{R}^2 |\ |xy| \leq c\}$$

$$Q_c = \{(x, y) \in \mathbb{R}^2 |\ |x| \leq \sqrt{c}\ e\ |y| \leq \sqrt{c}\}.$$

Claramente o quadrado Q_c está contido na região hiperbólica H_c, isto é, $Q_c \subset H_c$ para todo $c > 0$. O quadrado Q_c é limitado, simétrico



e convexo e sua área é $A(Q_c) = 4c$. Se tomarmos $c > A = 2\sqrt{d}$ (em que A representa a área de um domínio fundamental de L_1), estamos nas hipóteses do Teorema de Minkowski, Teorema 3.18, e assim, concluímos que existe um ponto não nulo do reticulado L_1 em Q_c. Ou seja, existe $\alpha_1 \in L_1$ tal que

$$0 \neq \alpha_1 = (a_1 + b_1\sqrt{d}, a_1 - b_1\sqrt{d}) \in Q_c \subset H_c \Rightarrow |a_1^2 - db_1^2| \leq c.$$

Podemos supor, por simetria, que $a_1, b_1 > 0$. Sejam $m = \min\{|a_1 + b_1\sqrt{d}|, |a_1 - b_1\sqrt{d}|\}$ (o menor dos módulos das coordenadas de α_1) e $\eta_2 \in]0, m\sqrt{\frac{d}{c}}[\subset \mathbb{R}$ ($\Rightarrow \eta_2 < \frac{c}{2}$).

Considere a transformação linear ortogonal:

$$T : \mathbb{R}^2 \to \mathbb{R}^2,\ T(x, y) = (\eta_2 x, \eta_2^{-1} y).$$

Afirmamos que a aplicação linear T transforma o reticulado L_1 num reticulado L_2 de mesma área (de um domínio fundamental). Com efeito, a transformação linear é bijetiva, logo a imagem do reticulado L_1 é também um reticulado L_2. Escolha η_2 tal que o reticulado L_2 não contenha o ponto α_1. Novamente, pelo teorema de Minkoswki, 3.18, existe um ponto não nulo do reticulado L_2 no quadrado Q_c,

$$0 \neq \alpha_2 = (a_2 + b_2\sqrt{d}, a_2 - b_2\sqrt{d}) \in Q_c \subset H_c \Rightarrow |a_2^2 - db_2^2| \leq c.$$

Podemos, assim, escolher $a_2, b_2 > 0$ tais que $|a_2 - b_2\sqrt{d}| \neq |a_1 - b_1\sqrt{d}|$.

Indutivamente, podemos construir uma infinidade de elementos $z_n \in \mathbb{Z}[\sqrt{d}]$ com valores absolutos distintos e com norma algébrica limitada. Isso implica na existência de um inteiro positivo $m \leq c$ tal que a equação

$$x^2 - dy^2 = m$$

possui uma infinidade de soluções inteiras. □

Prosseguimos, agora, com a demonstração do teorema 5.6.

Demonstração: do Teorema 5.6

Primeiramente vamos mostrar a existência de uma solução inteira positiva. Pela proposição 5.7, existe $m > 0$ tal que a equação

$$x^2 - dy^2 = m$$

possui uma infinidade de soluções inteiras. Podemos, portanto, escolher duas soluções positivas (x_1, y_1) e (x_2, y_2) tais que $|x_1| \neq |x_2|$, $x_1 \equiv x_2 \pmod{m}$ e $y_1 \equiv y_2 \pmod{m}$ (o número de classes de equivalência módulo m é finito ($= m$) portanto existem uma classe que contém uma infinidade de elementos).

Fazendo

$$(x_1 + y_1\sqrt{d})(x_2 - y_2\sqrt{d}) = (x_1x_2 - dy_1y_2) + (x_2y_1 - x_1y_2)\sqrt{d} \quad (5.1)$$

note que $x_1x_2 - dy_1y_2 \equiv x_1^2 - dy_1^2 \equiv 0 \pmod{m}$ e $x_1y_2 - x_2y_1 \equiv x_1y_1 - x_1y_1 \equiv 0 \pmod{m}$ pela nossa escolha de represantes na mesma classe de equivalência módulo m. Denotamos $x_1x_2 - dy_1y_2 = mu$ e $x_1y_2 - x_2y_1 = mv$. Substituindo na expressão 5.1, temos:

$$(x_1 + y_1\sqrt{d})(x_2 - y_2\sqrt{d}) = m(u + v\sqrt{d})$$

tomando conjugados:

$$(x_1 - y_1\sqrt{d})(x_2 + y_2\sqrt{d}) = m(u - v\sqrt{d})$$

multiplicando:

$$m^2(u^2 - dv^2) = (x_1^2 - y_1^2\sqrt{d})(x_2^2 - y_2^2\sqrt{d}) = m^2$$

Logo $u^2 - dv^2 = 1$. E assim mostramos que existe uma solução em inteiros positivos.

Para concluir devemos mostrar que todas as soluções em inteiros positivos são obtidas a partir de potências da solução mínima $\alpha = x_1 + y_1\sqrt{d}$ (claramente as potências desta solução sao também soluções, pela proposição 4.6). Suponha, por absurdo que exista uma solução positiva (x, y) que não pode ser obtida por potência, ou seja, $x + y\sqrt{d} \neq \alpha^n$, com $n \in \mathbb{N}$. Então existe um natural n tal que

$$\alpha^n < x + y\sqrt{d} < \alpha^{n+1}$$

isso implica

$$1 < \alpha^{-n}(x + y\sqrt{d}) < \alpha.$$

Isso é um absurdo pois $\alpha^{-n}(x + y\sqrt{d})$ é solução em inteiros positivos (Verifique!!!) e α é a solução mínima em inteiros positivos. □

5.4 Problemas

1. Encontre todas as soluções para a equação de Pell-Fermat $x^2 - dy^2 = 1$ nos casos em que $d = 3, 5, 6, 7, 8, 10$.

2. Mostre que se $d = c^2$, com $c \in \mathbb{N}$, ou seja, d é um quadrado perfeito. Então a equação $x^2 - dy^2 = m$ possui, sempre um número finito de soluções (pode ocorrer de não haver solução inteira). Encontre valores de d e de m para os quais a equação $x^2 - dy^2 = m$ possui soluções e valores para os quais a mesma não possui solução inteira positiva. Se $m = 1$ as únicas soluções são as triviais $(\pm 1, 0)$.

3. Mostre que as soluções inteiras positivas da equação $x^2 - 2y^2 = 1$ satisfazem a seguinte relação de recorrência: $(x_1, y_1) = (3, 2)$ e $x_{n+1} = 3x_n + 4y_n$, $y_{n+1} = 2x_n + 3y_n$.

4. Mostre que existem valores de d, não quadrados, para os quais a equação

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

possui solução e outros valores para os quais a mesma não admite solução.

5. Mostre que se (x_1, y_1) é a menor solução em inteiros positivos da equação $x^2 - dy^2 = -1$, então (x_2, y_2) definidos por $(x_2 + \sqrt{d}y_2) = (x_1 + \sqrt{d}y_1)^2$ é a menor solução em inteiros positivos da $x^2 - dy^2 = 1$.

6. Mostre que as soluções das equações $x^2 - dy^2 = \pm 1$ fornecem boa aproximações racionais para $\sqrt{d} \simeq \frac{x}{y}$. Sugestão: calcular $\frac{x}{y} - \sqrt{d}$

7. Sejam $d, n \in \mathbb{Z}$, com $d > 0$ não quadrado. Suponha que a hipérbole
$$x^2 - dy^2 = n$$
possua um ponto inteiro. Mostre que a mesma possui uma infinidade de pontos inteiros.

Bibliografias

[1] STEWART, I. N.; TALL, D. O. - *Algebraic number theory and Fermat's last Theorem*, A K Peters, Natick MA, 3rd Edition, 2002.

[2] CASSELS, J. W. S. - *Lectures on Elliptic Curves*, London Mathematical Society Student Texts, 24, 1991.

[3] GAMZON, A. - *The Hasse-Minkowski Theorem*, Honors Scholar Theses, University of Connecticut, 2006.

[4] BOREVICH, Z. I.; SHAFAREVICH, I. R. - *Number Theory*, Academic Press INc, Orlando Fl, 1966.

[5] GARCIA, A.; LEQUAIN, I. - *Elementos de Álgebra*, Projeto Euclides, IMPA, Rio de Janeiro, 1996.

[6] HEFEZ, A. - *Iniciação à Aritmética*, PIC-OBMEP, Rio de Janeiro, 2010.

[7] MIALARET, M. - *O Princípio Local-Global*, Monografia de Conclusão de Curso, UFRPE, 2008.

[8] GONDIM, R. - *A Aritmética das Curvas Algébricas Planas de Gênero Zero*, Notas de Aula, 2007.

[9] LAGES LIMA, E. - Meu Professor de Matemátrica e outras Histórias, Coleção do Professor de Matemática, IMPA, Riode Janeiro, 2006.

[10] HITCHIN, N. - Projective Geometry, b3 course 2003 (Lecture Notes), Oxford, 2003.

[11] MUNIZ NETO, A. C. - Equações Diofantinas, Revista Eureka! n. 7, Brasil, 2000.

[12] KUMAR DUTTA, A. - Mathematics in Ancient India-Diophantine Equations: The Kuttaka Resonance, An overview, Part 1, Vol. 7, No. 4, pp.4;19, Koltaka-India, 2002.

[13] WALDSCHIMIDT, M. - On the so called Pell-Fermat Equation www.impa.br/opencms/en, Rio de Janeiro, 2010.