Uma introdução à Geometria Algébrica Real

Caio De Naday Hornhardt

Dissertação apresentada
ao
Instituto de Matemática e Estatística
da
Universidade de São Paulo
para
obtenção do título
de
Mestre em Ciências

Programa: Matemática

Orientador: Prof. Dr. Fracisco Miraglia Neto

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro do CNPq

São Paulo, janeiro de 2014

Uma introdução à Geometria Algébrica Real

Esta é a versão original da dissertação elaborada pelo
candidato Caio De Naday Hornhardt, tal como
submetida à Comissão Julgadora.

Sumário

Agradecimentos

Muitos foram os amigos, familiares e profissionais que me ajudaram durante a realização deste trabalho.

Começo agradecendo minha namorada e eterna companheira Helen Samara dos Santos, que esteve do meu lado tanto nos melhores momentos quanto nos mais difíceis. Não consigo imaginar como eu teria feito esse trabalho sem ela.

Agradeço imensamente meus amigos Arthur Gabriel de Santana e Lucas Mendes Marques Gonçalves, os quais tiveram profundo impacto na minha forma de pensar e ver o mundo, o que certamente se reflete no texto que se segue. Fortaleci amizades com Vinícius Franco Vasconcelos e Pedro Vanalli Quirino mais próximo do fim desse trabalho, mas a importância deles na reta final foi grande.

A comunidade IME merece estar citada aqui, como um todo. Fui recebido e fiz parte dessa sociedade com muito orgulho nos últimos anos. Agradeço aos diversos servidores técnico-administrativos que tão bem me trataram e garantiram a infraestrutura necessária para meus estudos. Entre os docentes, além do grande apoio do meu orientador, Francisco Miraglia Neto, quero agradecer a dois em especial: Odilon Otávio Luciano e Paulo Agozzini Martin, com os quais tive iluminadores conversas e sugestões.

Outra pessoa de grande impacto e que não pode ser deixada de lado nessa seção é a Bruna Garcia Forlim. Agradeço profundamente a seu empenho e apoio, que tantos reflexos tiveram na minha vida.

Por fim, mas nem de longe menos importantes, agradeço a meus pais, Marta De Naday e Nelson Hornhardt, e à minha irmã, Rebeca De Naday Hornhardt, que acompanharam essa jornada desde o início com muito amor e suporte.

Resumo

HORNHARDT, Caio De N. **Uma Introdução à Geometria Algébrica**. 2014. 55 f. Dissertação (Mestrado) - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2014.

O presente trabalho traz uma apresentação geral da disciplina "Geometria Algébrica" e depois apresenta ferramentas para lidar com o caso de quando o corpo base é o corpo dos números reais. Os conceitos gerais, como variedades algébricas, topologia de Zariski e anéis de coordenadas, são introduzidos sobre corpos quaisquer e, posteriormente, são enfatizados os aspectos que envolvem a relação de ordem, que tem um papel fundamental quando estamos lidando com o corpo dos números reais. Apresentamos os conceitos de corpo formalmente real e corpo real fechado e suas diversas caracterizações, e provamos resultados clássicos do cálculo diferencial em uma variável para polinômios com coeficientes em um corpo real fechado. Essas ferramentas do cálculo são fundamentais para provarmos o Teorema de Tarski-Seidenberg, o qual é a peça chave da teoria. Esse teorema tem uma forte conexão com a Teoria de Primeira Ordem dos corpos reais fechados (e, portanto, do corpo dos números reais) e nossa prova deste é feita combinando os resultados do cálculo com teoremas clássicos da Teoria dos Modelos.

Palavras-chave: Geometria Algébrica, Teoria dos Aodelos, Álgebra Real

Abstract

HORNHARDT, Caio De N. **An Introduction to Real Algebraic Geometry**. 2014. 55 f. Dissertation (Master of Sciences) - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2014.

The present work shows a general presentation of the discipline called "Algebraic Geometry" and then presents tools for dealing with the case when the work is done above the field of the real numbers. General concepts, as algebraic varieties, Zariski topology and rings of coordinates, are introduced over any field and, after that, the aspects concerning the order relation, fundamental when we are working with the field of the real numbers, are emphasized. We present the concepts of formally real field and real closed field and their many characterizations and give proofs of classical results from one variable differential calculus for polynomials with coefficients in a real closed field. These tools from calculus are fundamental to our proof of the Tarski-Seidenberg theorem, with is a key result in the theory. This theorem has a strong connection with the first order theory of the real closed fields (and, therefore, of the field of the real numbers) and our proof combines the results of calculus with classical theorems of Model Theory.

Keywords: Algebraic Geometry, Model Theory, Real Algebra

Capítulo 1

Um Básico sobre Geometria Algébrica

1.1 Introdução

Uma resposta rápida para o que é *Geometria Algébrica* é o estudo de funções polinomiais (em várias variáveis), usando diversos domínios e contradomínios. O que exatamente quer dizer "diversos domínios e contradomínios" evoluiu bastante com o desenvolvimento da disciplina, mas a resposta mais natural, e historicamente a primeira, são os *subconjuntos* de k^n, onde k é o corpo dos coeficientes dos polinômios em questão.

Exemplo 1.1.1. Sejam $D = \{(x, y) \in \mathbb{R}^2 \mid y^2 = x \text{ e } x > 0\}$ e $F : D \longrightarrow \mathbb{R}$ tal que $F(x, y) = xy$

Mas geralmente, nosso objeto de estudo será:

Definição 1.1.2. Sejam $X \subseteq k^n$ e $Y \subseteq k^m$. Uma função $F : X \longrightarrow Y$ é dita *função polinomial* ou *morfismo afim* se existem polinômios $f_1, \ldots, f_m \in k[x_1, \ldots, x_n]$ tais que, para todo $P \in X$, $F(P) = (f_1(P), \ldots, f_m(P))$.

Observação. No decorrer do capítulo, k é um corpo e n, m são números naturais fixados.

Com a função descrita em 1.1.1, podemos exemplificar os dois primeiros conceitos que vamos abordar nesse capítulo. Primeiramente, que o *polinômio* que define a função, em geral, não é único. A mesma F pode ser escrita como $F(x, y) = x^3$, por causa do domínio em que está definida. É claro que esta função pode ser estendida para todo o plano, polinomialmente. Basta definir nos pontos restantes como $F_1(x, y) = xy$ ou $F_2(x, y) = x^3$, mas estas extensões são *distintas no plano.*

Entretanto, existem subconjuntos de $\mathbb{R}^2$ onde esta função pode ser estendida *de maneira unica.* Por exemplo, usando a topologia usual em $\mathbb{R}$ e $\mathbb{R}^2$, como F tem extensão contínua e 0 é ponto de acumulação de D, temos que

$$\lim_{(x,y)\to(0,0)} g(x, y) = \lim_{(x,y)\to(0,0)} xy$$

qualquer que seja o polinômio $g \in \mathbb{R}[x, y]$ contando que represente a mesma função que F quando restrito a D.

Na verdade, podemos estender para um subconjunto muito maior.

Definição 1.1.3 (fecho de Zariski)**.** Seja X um subconjunto de k^n. Chamamos de *fecho de Zariski* de X o conjunto

$$\overline{X} = \{P \in k^n \mid f(P) = g(P) \text{ para todos os } f, g \in k[x_1, \ldots, x_n] \text{ tais que } f \restriction_X = g \restriction_X\}.$$

O fecho de Zariski é o maior domínio para uma extensão que não altera as possíveis "fórmulas" escolhíveis para representar uma função; e esse fecho define uma topologia conveniente para lidar com polinômios em k^n, conhecida como *topologia de Zariski*.

Para verificar que a operação definida em 1.1.3 é, de fato, uma *operação de fecho* em um espaço topológico, vamos verificar os axiomas de Kuratowski para este tipo de construção (o leitor que desconhece essa maneira de definir um espaço topológico pode consultar [4])

Mas antes disso, vamos dar uma pequena simplificada na definição:

Proposição 1.1.4. *Seja $X \subseteq k^n$. Então*

$$\overline{X} = \{P \in k^n \mid f(P) = 0 \text{ para todo polinômio } f \in k[x_1, \ldots, x_n] \text{ tal que } f \restriction_X = 0\}$$

Demonstração. É claro que o conjunto do lado direito da igualdade no enunciado está contido em $\overline{X}$. Para a outra inclusão, basta notar que se $P \in k^n$ e $f, g \in k[x_1, \ldots, x_n]$, então

$$f \restriction_X = g \restriction_X \iff (f - g) \restriction_X = 0. \qquad \blacksquare$$

Proposição 1.1.5 (topologia de Zariski)**.** *Sejam $X, Y \subseteq k^n$. A operação fecho definida acima satisfaz:*

$$(i)\ X \subseteq \overline{X}; \qquad (ii)\ \overline{\overline{X}} = \overline{X}; \qquad (iii)\ \overline{X \cup Y} = \overline{X} \cup \overline{Y}; \qquad (iv)\ \overline{\emptyset} = \emptyset,$$

constituindo-se, portanto, em uma operação de fecho de um espaço topológico. A essa topologia é dado o nome de Topologia de Zariski.

Demonstração. O item (i) é imediato da definição.

(ii) Por (i), basta provar que $\overline{\overline{X}} \subseteq \overline{X}$. Seja $P \in \overline{\overline{X}}$ e seja $f \in k[x_1, \ldots, x_n]$ tal que f coincide com a função nula sobre X. Então, por definição de fecho, f coincide com a nula em $\overline{X}$, de onde concluímos que $f(P)$ também tem que valer 0.

(iii) Seja $P \in \overline{X} \cup \overline{Y}$. Sem perda de generalidade, podemos supor $P \in \overline{X}$. Se $f \in k[x_1, \ldots, x_n]$ se anula sobre $X \cup Y$, em particular de anula sobre X, de onde concluímos que $f(P) = 0$.

Para a outra inclusão, seja $Q \in k^n$ tal que $Q \notin \overline{X} \cup \overline{Y}$. Então existem $f, g \in k[x_1, \ldots, x_n]$ tais que f se anula em X, mas $f(Q) \neq 0$, e g se anula sobre Y, mas $g(Y) \neq 0$. O polinômio $f \cdot g$ se anula sobre a união $X \cup Y$, logo se anula em $\overline{X \cup Y}$. Como $f(Q) \cdot g(Q) \neq 0$, $Q \notin \overline{X \cup Y}$.

(iv) Quaisquer duas funções coincidem se restritas ao vazio; mas as constantes definidas pelos polinômios 1 e 0 diferem para qualquer ponto $P \in k^n$. $\blacksquare$

Definição 1.1.6. A um conjunto fechado na topologia de Zariski também são dados os nomes de *conjunto algébrico* ou de *variedade algébrica afim*.

Agora vamos verificar que com essa topologia as funções polinomiais são contínuas (usando a caracterização de continuidade por fechos topológicos):

Proposição 1.1.7. *Sejam $X \subseteq k^n$, $Y \subseteq k^m$ e $F : X \longrightarrow Y$ uma função polinomial. Para qualquer $S \subseteq X$, temos que $F(\overline{S}) \subseteq \overline{F(S)}$ (os fechos são relativos aos subespaços X e Y).*

Demonstração. Suponhamos que existam $P \in \overline{S}$ e um polinômio $g \in k[x_1, \ldots, x_m]$ tais que $g \restriction_{F(S)} = 0$ mas $g(F(P)) \neq 0$.

Sejam $f_1, \ldots, f_m \in k[x_1, \ldots, x_n]$ tais que $F = (f_1, \ldots, f_m)$ e escrevemos

$$g = \sum a_{\alpha_1, \ldots, \alpha_m} x_1^{\alpha_1} \cdots x_m^{\alpha_m},$$

com os $\alpha_1, \ldots, \alpha_m$ variando sobre $\mathbb{N}$. Assim, a função $g \circ F$ é polinomial e dada pelo polinômio $g* = \sum a_{\alpha_1,\ldots,\alpha_m} f_1^{\alpha_1} \cdots f_m^{\alpha_m} \in k[x_1, \ldots, x_n]$. Como $g \circ F = 0$, $g* \restriction_S = 0$, logo $g*(P) = g(F(P)) = 0$, um absurdo. ■

1.1.1 A Correspondência $\mathcal{Z}$

Para lidarmos melhor com a topologia de Zariski, vamos fazer uma nova caracterização dos conjuntos fechados. Para tal, daremos a seguinte definição:

Definição 1.1.8 (correspondência $\mathcal{Z}$)**.** Para todo $S \subseteq k[x_1, \ldots, x_n]$, definimos $\mathcal{Z}(S)$ como sendo o seguinte subconjunto de k^n:

$$\mathcal{Z}(S) \doteq \{P \in k^n \mid f(P) = 0, \forall f \in S\}.$$

1.2 O Anel de Coordenadas

1.2.1 A Correspondência $\mathcal{I}$

A correspondência $\mathcal{Z}$ possui uma espécie de inversa, que chamaremos de $\mathcal{I}$:

Definição 1.2.1 (correspondência $\mathcal{I}$)**.** Para todo $X \subseteq k^n$, definimos $\mathcal{I}(X)$ como sendo o seguinte subconjunto de $k[x_1, \ldots, x_n]$:

$$\mathcal{I}(X) \doteq \{f \in k[x_1, \ldots, x_n] \mid f \restriction_X = 0\}.$$

A função $\mathcal{I}$ é uma boa medida para o quanto a representação de uma função polinomial não é única. Como já dissemos, os polinômios f e g representam a mesma função sobre um conjunto X se, e somente se, $f - g$ representa a função nula em X, isto é, $f - g \in \mathcal{I}(X)$. É de se esperar que $\mathcal{I}$ tenha um papel central no estudo das funções polinomiais sobre um dado conjunto X.

Notação 1.2.2. Sejam $X \subseteq k^n$ e $Y \subseteq k^m$. Denotamos o *conjunto das funções polinomiais de X em Y* por $\mathcal{F}(X, Y)$.

A próxima afirmação decorre facilmente da definição de função polinomial e da definição de $\mathcal{I}(X)$, por isso sua demonstração será omitida.

Proposição 1.2.3. *A função $\phi : k[x_1, \ldots, x_n] \longrightarrow \mathcal{F}(X, k)$ dado por $\phi(f)(P) = f(P)$ é um homomorfismo sobrejetor, com $ker\phi = \mathcal{I}(X)$. Em particular $\mathcal{I}(X)$ é um ideal e $\frac{k[x_1,\ldots,x_n]}{\mathcal{I}(X)} \cong \mathcal{F}(X, k)$.*

Definição 1.2.4. Dado $X \subseteq k^n$, damos o nome de *anel das coordenadas de X* ao anel $\frac{k[x_1,\ldots,x_n]}{\mathcal{I}(X)}$, e o denotamos por $k[X]$. Identificaremos $k[X]$ com $\mathcal{F}(X, k)$ pela proposição acima.

Quando não houver perigo de confusão, denotaremos um elemento de $k[X]$ por qualquer representante de sua classe de equivalência em $k[x_1, \ldots, x_n]$.

Observação. O anel $k[X]$ é também um espaço vetorial sobre k, usando a mesma soma e restringindo a multiplicação da maneira conveniente, *i.e.*, $k[X]$ é uma k-álgebra, *i.e.*, um anel que também é uma espaço vetorial sobre k. Como k-álgebra, $k[X]$ é *finitamente gerada*.

Essa definição é simplesmente um nome para a identificação natural dada em 1.2.3. O nome "coordenadas" se deve ao fato que este é o menor anel que contém as "funções coordenadas", isto é, as classes de equivalência dos polinômios $x_1, \ldots, x_n$.

1.2.2 Funções Polinomiais e Homomorfismos

Uma coisa bastante interessante sobre os anéis de coordenadas é que os homomorfismos entre eles correspondem a funções polinomiais entre os conjuntos subjacentes (se estes forem algébricos) e vice-versa.

Proposição 1.2.5. *Sejam $X \subseteq k^n$, $Y \subseteq k^m$ e $f \in \mathcal{F}(X,Y)$. Então temos que $\Phi_f : k[Y] \longrightarrow k[X]$ dada por $\psi_f(g) = g \circ f$ é um homomorfismo k-álgebras (*i.e.*, um homomorfismo de anéis que também é k-linear).*

Demonstração. Sejam $a, b \in k[Y]$, $\lambda \in k$ e $P \in X$. Basta fazer as contas seguindo as definições das operações com funções:

$$(a + \lambda \cdot b)(f(P)) = a(P) + \lambda \cdot b(P) \quad \text{e} \quad (a \cdot b)(f(P)) = a(f(P)) \cdot b(f(P)). \qquad \blacksquare$$

Proposição 1.2.6. *Se Y é* algébrico*, dado um homomorfismo $\varphi : k[Y] \longrightarrow k[X]$, temos que $F_\phi : X \longrightarrow Y$, dada por $F_\varphi(P) = (\varphi(x_1)(P), \ldots, \varphi(x_m)(P))$, é uma função polinomial.*

Demonstração. Para uma fórmula para F_φ com polinômios, basta pegar um representante em $g_i \in k[x_1, \ldots, x_n]$ para cada $\varphi(x_i)$. A única coisa que de fato precisa ser mostrada é que podemos tomar Y como contradomínio, isto é, que

$$(\varphi(x_1)(P), \ldots, \varphi(x_m)(P)) = (g_1(P), \ldots, g_m(P)) \in Y,$$

para todo $P \in X$. Para isso vamos usar que Y é fechado na topologia de Zariski. Pela prova da proposição 1.1.5, temos que mostrar que se $f \in \mathcal{I}(Y)$, então $f(g_1(P), \ldots, g_m(P)) = 0$. Definimos

$$f = \textstyle\sum_{\alpha \in \mathbb{N}^m} a_\alpha \cdot x_1^{\alpha_1} \cdots x_m^{\alpha_m},$$

onde α_i é a i-ésima coordenada de α e $a_\alpha \in k$, para todo α. Então

$$f(g_1(P), \ldots, g_m(P)) = \textstyle\sum_{\alpha \in \mathbb{N}^m} a_\alpha \cdot g_1(P)^{\alpha_1} \cdots g_m(P)^{\alpha_m} = h(P),$$

onde $h(x_1, \ldots, x_n) = \sum_{\alpha \in \mathbb{N}^m} a_\alpha \cdot g_1^{\alpha_1} \cdots g_m^{\alpha_m}$. Basta mostrar que $h \in \mathcal{I}(X)$, pois então $f(g_1(P), \ldots, g_m(P)) = h(P) = 0$, se $P \in X$. Com efeito, em $k[X]$, temos, pela escolha dos g_i, que

$$h = \textstyle\sum_{\alpha \in \mathbb{N}^m} a_\alpha \cdot (\varphi(x_1))^{\alpha_1} \cdots (\varphi(x_m))^{\alpha_m},$$

que é igual, por φ ser homomorfismo, à $\varphi(\sum_{\alpha \in \mathbb{N}^m} a_\alpha \cdot x_1^{\alpha_1} \cdots x_m^{\alpha_m}) = \varphi(f)$, onde f aqui representa a classe de f em $k[Y]$. Como pegamos $f \in \mathcal{I}(Y)$, $h = \varphi(0) = 0$ em $k[X]$, logo $h \in \mathcal{I}(X)$. $\blacksquare$

Notação 1.2.7. a) Sejam A um anel commutativo com unidade e $S \subseteq A$. Escrevemos:

(*i*) $S \lhd A$ para indicar que S é um *ideal* em A;

(*ii*) $\langle S \rangle$ para indicar o *ideal gerado* por S em A.

b) Sejam A e B k-álgebras. Denotamos o conjunto dos homomorfismos de k-álgebras entre A em B por $Hom(A, B)$.

Proposição 1.2.8. *Sendo Y algébrico, as funções $\Phi : \mathcal{F}(X,Y) \longrightarrow Hom(k[Y], k[X])$ e $F : Hom(k[Y], k[X]) \longrightarrow \mathcal{F}(X,Y)$, definidas acima, são inversas.*

Demonstração. Vamos calcular $F \circ \Phi$. Seja $f \in \mathcal{F}(X,Y)$, $f = (f_1, \ldots, f_m)$ com $f_1, \ldots, f_m \in k[x_1, \ldots, x_n]$. Queremos mostrar que $(F \circ \Phi)(f)$ coincide com f em todo ponto $P \in X$. Com efeito,

$$\begin{aligned}(F \circ \Phi)(f)(P) &= F_{\Phi_f}(P) = (\Phi_f(x_1)(P), \ldots, \Phi_f(x_m)(P)) \\ &= ((x_1 \circ f)(P), \ldots, (x_m \circ f)(P)) = (f_1(P), \ldots, f_m(P)) = f(P).\end{aligned}$$

Agora vamos calcular $\Phi \circ F$. Seja $\varphi \in Hom(k[Y], k[X])$. Queremos mostrar que $(\Phi \circ F)(\varphi)$ coincide com φ para todo elemento de $k[Y]$. Como ambas são homomorfismos de k-álgebras, basta verificar que coincidem nos elementos $x_1, \ldots, x_m \in k[Y]$. Seja $i \in \{1, \ldots, m\}$. Então

$$(\Phi \circ F)(\varphi)(x_i) = \Phi_{F_\varphi}(x_i) = x_i(F_\varphi),$$

isto é, a i-ésima coordenada de F_φ. Mas $F_\varphi(P) = (\varphi(x_1)(P), \ldots, \varphi(x_m)(P))$, cuja i-ésima coordenada é $\varphi(x_i)(P)$. Pela arbitrariedade de $P \in X$, temos a igualdade desejada. ■

Esse resultado mostra uma importante conexão[1] da teoria dos anéis comutativos com unidade (chamada de *Álgebra Comutativa*) com a Geometria Algébrica. Não são todos as k-álgebras finitamente geradas que são da forma $k[X]$ para algum X, mas certamente algum conhecimento sobre estas será útil. Boa parte dos conceitos da Álgebra Comutativa, como localização ou dimensão de Krull, usados em casos que vão além dos anéis dessa forma, são inspirados pela Geometria Algébrica.[2]

1.3 Um pouco mais sobre as funções $\mathcal{Z}$ e $\mathcal{I}$

As funções $\mathcal{Z}$ e $\mathcal{I}$ não são, de fato, inversas, mas se aproximam disso. Elas se tornam inversas uma da outra quando restritas "aos conjuntos que mais interessam". Como corolário da demonstração de 1.1.5, é possível notar que $\mathcal{Z}(\mathcal{I}(X)) = X$ se, e somente se, X é algébrico. Vamos investigar um pouco mais essas funções e a relação entre elas.

Proposição 1.3.1. *Sejam $X \subseteq k^n$ e $S \subseteq k[x_1, \ldots, x_n]$. Então*

$$X \subseteq \mathcal{Z}(S) \text{ se, e somente se, } S \subseteq \mathcal{I}(X).$$

Demonstração. Seja $f \in S$ e suponhamos $X \subseteq \mathcal{Z}(S)$, então $f \restriction_X = 0$, ou seja, $f \in \mathcal{I}(X)$. Analogamente, seja $x \in X$ e suponhamos $S \subseteq \mathcal{I}(X)$. Isso significa que $f(x) = 0$ para todo $f \in S$, ou seja, $x \in \mathcal{Z}(S)$. ■

Vamos ver agora algumas consequências básicas da proposição acima

Corolário 1.3.2. *Com a notação acima, temos:*

(i) As compostas $\mathcal{Z} \circ \mathcal{I}$ e $\mathcal{I} \circ \mathcal{Z}$ são inflacionárias *(*i.e.*, funções f tais que $x \leq f(x)$).*

*(ii) As compostas $\mathcal{Z} \circ \mathcal{I}$ e $\mathcal{I} \circ \mathcal{Z}$ são idempotentes (*i.e.*, funções f tais que $f \circ f = f$).*

(iii) As funções $\mathcal{Z}$ e $\mathcal{I}$ são decrescentes.

(iv) $(\mathcal{Z} \circ \mathcal{I} \circ \mathcal{Z})(S) = \mathcal{Z}(S)$ *e* $(\mathcal{I} \circ \mathcal{Z} \circ \mathcal{I})(X) = \mathcal{I}(X)$.

Demonstração. Por simetria, vamos apenas mostrar que $\mathcal{I} \circ \mathcal{Z}$ é inflacionária e idempotente, que $\mathcal{Z}$ é decrescente e que $(\mathcal{Z} \circ \mathcal{I} \circ \mathcal{Z})(S) = \mathcal{Z}(S)$.

(i) Como $\mathcal{Z}(S) \subseteq \mathcal{Z}(S)$, pela propriedade da proposição 1.3.1, $S \subseteq \mathcal{I}(\mathcal{Z}(S))$.

(ii) Pelo item anterior, temos que $(\mathcal{I} \circ \mathcal{Z})(S) \subseteq (\mathcal{I} \circ \mathcal{Z}) \circ (\mathcal{I} \circ \mathcal{Z})(S)$. A outra inclusão decorre dessa, usando a propriedade da proposição 1.3.1.

(iii) Sejam $S \subseteq T \subseteq k[x_1, \ldots, x_n]$. Então $S \subseteq \mathcal{I}(\mathcal{Z}(T))$, de onde $\mathcal{Z}(S) \supseteq \mathcal{Z}(T)$.

(iv) Já sabemos que $\mathcal{Z}(S) \subseteq (\mathcal{Z} \circ \mathcal{I} \circ \mathcal{Z})(S)$. Para a outra inclusão, basta usar que $(\mathcal{I} \circ \mathcal{Z})(S) \subseteq (\mathcal{I} \circ \mathcal{Z})(S)$ e a propriedade da proposição 1.3.1. ■

[1] Ao leitor que conhece um pouco de teoria das categorias, devo ressaltar que este resultado pode ser, sem grande esforço, melhorado para uma equivalência (contravariante) entre duas categorias

[2] De fato, é possível fazer uma teoria para "fingir" que um anel comutativo qualquer é um "espaço de funções", com os conceitos de *espectro de um anel* e de *esquema*.

Corolário 1.3.3. *As funções $\mathcal{I}$ e $\mathcal{Z}$ são inversas quando restritas uma à imagem da outra.*

Por hora, vamos nos restringir aos subconjuntos de $k[x_1, \ldots, x_n]$ que são *ideais*. O corolário acima garante que não perdemos nada com isso. Ainda resta a pergunta se não podemos nos restringir *ainda mais*, isto é, se a imagem de $\mathcal{I}$ contém todos ideais ou só alguns. Este será o tema da próxima seção.

Proposição 1.3.4. *Os zeros simultâneos de $S \subseteq k[x_1, \ldots, x_n]$ coincidem com os do ideal gerado por S,*i.e.*, $\mathcal{Z}(S) = \mathcal{Z}(\langle S \rangle)$.*

Demonstração. Saí direto da definição de ideal, ou ainda do fato que $S \subseteq \langle S \rangle \subseteq (\mathcal{I} \circ \mathcal{Z})(S)$ e aplicando $\mathcal{Z}$ a todos os termos nessa desigualdade. ■

O momento parece conveniente para o seguinte lema, que usaremos posteriormente:

Lema 1.3.5. *Sejam $I, J \triangleleft k[x_1, \ldots, x_n]$. Então $\mathcal{Z}(I \cap J) = \mathcal{Z}(I) \cup \mathcal{Z}(J)$.*

Demonstração. Seja $x \in \mathcal{Z}(I) \cup \mathcal{Z}(J)$ e $f \in I \cap J$. O ponto x é zerado por todo elemento de I ou por todo elemento de J. Em ambos casos, $f(x) = 0$.

Se $x \notin \mathcal{Z}(I) \cup \mathcal{Z}(J)$, então existem $f \in I$ e $g \in J$ tais que $f(x) \neq 0 \neq g(x)$. Então $f \cdot g(x) \neq 0$, mas $f \cdot g \in I \cap J$, de onde $x \notin I \cap J$. ■

1.3.1 Os ideais da forma $\mathcal{I}(X)$

Agora nos deparamos com um problema importante. De um lado do isomorfismo 1.3.3 temos os conjuntos algébricos. Naturalmente, queremos saber melhor quem está do outro lado, os ideais que são $\mathcal{I}(X)$ para algum X. Já sabemos que estes determinam completamente os anéis $k[X]$.

Uma primeira pergunta seria se todos os ideais são dessa forma. A resposta é negativa.

Exemplo 1.3.6. Tome $I \doteq \langle x_1^3 \rangle \triangleleft k[x_1, \ldots, x_n]$. Um ponto $(p_1, \ldots, p_n)$ está em $\mathcal{Z}(I)$ se, somente se, $p_1^3 = 0$ se, e somente se, $p_1 = 0$. Logo $I = \langle x_1^3 \rangle \subsetneq \langle x_1 \rangle \subseteq \mathcal{I}(\mathcal{Z}(I))$. Mas então I não pode estar na imagem de $\mathcal{I}$, pelo corolário 1.3.2.

Esse exemplo nos dá uma ideia de qual é o próximo passo que devemos dar:

Definição 1.3.7. Sejam A um anel e $I \triangleleft A$. Chamamos de *radical de I* ao conjunto

$$\sqrt{I} \doteq \{f \in A \mid \exists k \in \mathbb{N},\ f^k \in I\}.$$

Definição 1.3.8. Um ideal $I \triangleleft A$ é chamado de *ideal radical* se $\sqrt{I} = I$.

Proposição 1.3.9. *O radical de um ideal $I \triangleleft A$ é um ideal. Mais que isso, é um ideal radical.*

Demonstração. Sejam $a, b \in \sqrt{I}$ e $c \in A$, e $l, m \in \mathbb{N}$ tais que $a^l, b^m \in I$. $(ca)^l = c^l a^l \in I$. Também temos que

$$(a+b)^{l+m} = \sum_{i=0}^{l+m} \binom{l+m}{i} a^{l+m-i} b^i \in I$$

pois cada termo $a^{l+m-i} b^i \in I$ (Se $i \leq m$, então $a^{l+m-i} \in I$. Se $i \geq m$, então $b^i \in I$).

Com isso temos que $\sqrt{I}$ é um ideal. Que é um ideal radical é fácil: se $a^r \in \sqrt{I}$ então existe $k \in \mathbb{N}$ com $(a^r)^k = a^{rk} \in I$, logo $a \in \sqrt{I}$. ■

Proposição 1.3.10. *Se $X \subseteq k^n$, então $\mathcal{I}(X)$ é um ideal radical.*

Demonstração. Se $f^l \in \mathcal{I}(X)$, então $f^l(x) = 0$ para todo $x \in X$, mas como os valores estão em um corpo, $f(x) = 0$ para todo $x \in X$, isto é, $f \in \mathcal{I}(X)$. ■

Os ideais radicais são suficientes?

Agora podemos refinar nossa pergunta: os ideais da forma $\mathcal{I}(X)$ são todos os ideais radicais?

A resposta é a seguinte: depende.

O problema é que não é possível classificar esses ideais com toda a generalidade que tratamos até o momento. Para diferentes corpos temos diferentes respostas.

Para o caso dos corpos *algebricamente fechados*, Hilbert provou que nada mais precisa ser feito, os ideais radicais são exatamente os ideais na imagem de $\mathcal{I}$. Esse resultado ficou conhecido pelo nome em alemão, *Nullstellensatz*, e é a principal razão de porque a Geometria Algébrica se desenvolveu principalmente sobre corpos algebricamente fechados.

Para outros corpos precisamos tomar subclasses próprias dos ideais radicais. Podemos ver isso com o seguinte exemplo:

Exemplo 1.3.11. Em $\mathbb{R}^2$, tomemos $X = \mathcal{Z}(\langle x^2 + (y-1)^2 \rangle)$. Como se trata de uma soma de quadrados, temos que $X = \{(0,1)\}$ e portanto $x \in \mathcal{I}(X)$, mas $x^k \notin \langle x^2 + (y-1)^2 \rangle$, independentemente de $k \in \mathbb{N}$. Em particular, o ideal $\sqrt{\langle x^2 + (y-1)^2 \rangle}$ não está na imagem de $\mathcal{I}$.

É útil entender um pouco mais sobre os ideais radicais. Daremos alguns resultados sobre eles na seção 1.4.3. Mas antes, vamos olhar um pouco para os ideais de $k[x_1, \ldots, x_n]$ em geral.

1.4 Anéis e Espaços Topológicos Noetherianos

1.4.1 Anéis Noetherianos

Como podemos notar, os ideais de $k[x_1, \ldots, x_n]$ desempenham um papel fundamental. No caso dos polinômios em uma variável, muito coisa provém do fato de todos os ideais serem principais. Aqui não temos tanta sorte, mas conseguimos garantir que todo ideal é *finitamente gerado*.

Proposição 1.4.1. *Seja A um anel*

São equivalentes:

(*i*) *Todos os ideais de A são finitamente gerados;*

(*ii*) *Os ideais de A satisfazem a* Condição da Cadeia Ascendente (CCA)*, isto é, se*

$$I_1 \subseteq I_2 \subseteq \cdots \subseteq I_n \subseteq \cdots$$

é uma cadeia ascendente de ideais, então existe $N \in \mathbb{N}$ tal que $I_i = I_j$ para todo $i, j \geq N$;

(*iii*) *Todo conjunto não vazio de ideais de A possui um elemento maximal.*

Demonstração. (*i*) *implica* (*ii*)*:* Seja $\{I_k\}_{k\in\mathbb{N}}$ uma cadeia ascendente de ideais de A. É fácil ver que $\bigcup_{k\in\mathbb{N}} I_k$ é um ideal, o qual tem que ter um conjunto gerador finito, digamos $\{a_1, \ldots, a_l\}$. Cada a_i é encontrado em algum I_{k_i}, logo todos estão em I_N onde $N \doteq max\{k_1, \ldots, k_l\}$; assim, $I_N = \bigcup_{k\in\mathbb{N}} I_k = I_t$, para todo $t \geq N$.

(*ii*) *implica* (*iii*)*:* Seja $\mathcal{C}$ um conjunto não vazio de ideais de A *sem* elemento maximal. Seja I_1 um elemento em $\mathcal{C}$. Como I_1 não é maximal, existe $I_2 \in \mathcal{C}$ com $I_1 \subsetneq I_2$. Seguindo esse raciocínio, encontramos uma cadeia ascendente de ideais que não é estacionária.

(*iii*) *implica* (*i*)*:* Seja J um ideal de A e seja $\mathcal{C} = \{\langle S \rangle \mid S \subseteq J \text{ é finito. }\}$. Por hipótese, existe $J' \in \mathcal{C}$ maximal em $\mathcal{C}$. Sendo assim, para todo $a \in J$ temos que $\langle J' \cup \{a\} \rangle = J'$, de onde concluímos que $J' = J$. ∎

Definição 1.4.2 (anel noetheriano)**.** Um anel satisfazendo uma (e, portanto, todas) das afirmações acima é dito *noetheriano*.

Definição 1.4.3. Uma sequência de elementos $\{x_k\}_{k\in\mathbb{N}}$ de um conjunto qualquer é dita *estacionária* se existe $n \in \mathbb{N}$ tal que $x_i = x_N$ para todo $i \geq N$.

Teorema 1.4.4 (da Base de Hilbert)**.** *Se A é noetheriano e x é uma variável independente sobre A, então $A[x]$ é noetheriano.*

Demonstração. Seja $I \triangleleft A[x]$. Provemos que I é finitamente gerado. Para cada $n \in \mathbb{N}$, definamos

$$J_n = \{a \in A \mid \exists p(x) \in I \text{ tal que } p(x) = ax^n + g(x)$$
$$\text{e o grau de } g(x) \text{ é menor que } n \text{ ou } g(x) = 0\}$$

Cada J_n é um ideal de A, pois $0 \in J_n$ (basta tomar $p(x) = 0$), se $a, b \in J_n$ temos que $a + b \in J_n$ (basta somar os respectivos polinômios) e se $a \in J_n$ e $c \in A$, $ca \in J_n$ (bastando fazer $c \cdot p(x)$).

Como A é noetheriano, para cada J_n podemos tomar um conjunto gerador G_n finito e, para evitar burocracias, sem o 0. Além disso, os ideais da forma J_n são em quantidade finita: de fato, temos que $J_n \subseteq J_{n+1}$, pois se $p(x) = p(x) = ax^n + g(x)$, com grau $g < n$ ou $g = 0$, $x \cdot p(x) = ax^{n+1} + g_1(x)$, com grau $g_1(x) < n + 1$ ou $g_1(x) = 0$; assim existe $N \in \mathbb{N}$ tal que $J_n = J_N$, para todo $n \geq N$.

Sendo assim, também podemos supor que $G_n = G_N$, para todo $n \geq N$.

Vamos construir um conjunto gerador para I: se $n \leq N$, para cada $a \in G_n$ escolhermos um polinômio $p_{a,n}(x)$ de grau n com a sendo o coeficiente de seu termo de maior grau (lembremos que $0 \notin G_n$). Chamemos o conjunto de todos esses polinômios de G e provemos que $\langle G\rangle = I$.

Se $n > N$ e $a \in G_n = G_N$, chamemos de $p_{a,n}$ o polinômio $x^{n-N} \cdot p_{a,N}$. Claramente $p_{a,n} \in \langle G\rangle$.

Seja $f(x) \in I$. Se f é constante, $f \in J_0 = \langle G_0\rangle$ e $G_0 \subset G$. Suponhamos que já temos tal resultado para todas os polinômios de grau menor ou igual a um dado n e que o grau de f é $n+1$. Se $b \in A$ é o coeficiente do termo de maior grau de f, temos que

$$b = \textstyle\sum_{i=0}^{m} c_i \cdot a_i$$

com $a_i \in G_{n+1}$. Assim, o polinômio

$$g(x) \doteq f(x) - \textstyle\sum_{i=0}^{m} c_i \cdot p_{a_i,n+1}(x)$$

tem grau menor ou igual a n, logo $g \in \langle G\rangle$, pela hipótese de indução; como $\sum_{i=0}^{m} c_i p_{a_i}(x)$ obviamente está em $\langle G\rangle$, obtemos $f \in \langle G\rangle$, como desejado. ∎

Corolário 1.4.5. *Todo conjunto algébrico é da forma $\mathcal{Z}(S)$, onde $S \subseteq k[x_1, \ldots, x_n]$ é* finito*.*

Demonstração. Basta tomar $\mathcal{Z}$(soma de quadrados) ∎

Corolário 1.4.6. *Se $X \subseteq k^n$, então $k[X]$, o anel de coordenadas de X é noetheriano.*

Demonstração. Basta usar o teorema da correspondência de ideais para ver que todo quociente de um anel noetheriano é noetheriano. ∎

1.4.2 Espaços Topológicos Noetherianos

O teorema da base de Hilbert (1.4.4) nos dá informação sobre a topologia de Zariski.

Definição 1.4.7. Um espaço topológico é dito **noetheriano** se seus conjuntos *fechados* satisfazem a *Condição da Cadeia Descendente (CCD)*, isto é, se $\{F_k\}_{k\in\mathbb{N}}$ é uma sequência de fechados tal que

$$F_0 \supseteq F_1 \supseteq \cdots \supseteq F_l \supseteq \cdots,$$

então existe $N \in \mathbb{N}$ tal que se $l \geq N$ então $F_l = F_N$.

É claro que este é o caso de k^n (e portanto, de seus subespaços) com a topologia de Zariski, por causa de 1.3.2(*iii*).

Proposição 1.4.8. *Um espaço topológico X é noetheriano se, e somente se, todo subespaço de X é compacto.*

Demonstração. Seja $S \subseteq X$ e $\mathcal{C} = \{U_\lambda\}_{\lambda\in\Lambda}$ uma cobertura por abertos de S que *não* possui subcobertura finita. Construímos uma sequencia $\{U_i\}_{i\in\mathbb{N}}$ da seguinte maneira: tomo U_0 como sendo um aberto qualquer de $\mathcal{C}$ e U_{n+1} como sendo U_n unido com algum aberto que contém algum elemento de X não coberto por U_n. Assim temos que $U_0 \subsetneq U_1 \subsetneq \cdots U_i \subsetneq \cdots$, logo a sequência dos complementares de U_i é uma sequência decrescente de fechados que não satisfaz a CCD.

Para a recíproca, se X não é noetheriano, existe uma sequência de fechados $\{F_i\}_{i\in\mathbb{N}}$ tais que $F_0 \supsetneq F_1 \supsetneq \cdots F_n \supsetneq \cdots$. Para cada $i \in \mathbb{N}$, seja x_i um elemento que está em F_i mas não está em F_{i+1}. Afirmamos que $S \doteq \{x_i \mid i \in \mathbb{N}\}$ não é compacto. Se o fosse, $\{X\backslash F_i \mid i \in N\}$ seria uma cobertura por abertos para S, e portanto conteria uma subcobertura finita. Se $N \in \mathbb{N}$ é o maior índice de $X\backslash F_i$ que aparece numa subcobertura dessas, temos que $\{X\backslash F_N\}$ *per se* é uma subcobertura, o que é um absurdo pois $x_N \in F_n \cap S$. ■

1.4.3 Variedades Irredutíveis e Ideais Primos

Iremos usar a condição da cadeia descendente construir uma decomposição dos conjuntos algébricos análoga à fatoração em números primos em $\mathbb{Z}$.

Definição 1.4.9. Um espaço topológico noetheriano X é dito *irredutível* se não é possível escrever $X = A \cup B$ com A, B dois conjuntos *fechados* distintos de X.

Proposição 1.4.10. *Um subconjunto $X \subseteq k^n$ com a topologia de Zariski é irredutível se, e somente se, $\mathcal{I}(X)$ é um ideal primo.*

Demonstração. Suponhamos que $\mathcal{I}(X)$ não é primo, e sejam $f, g \in k[x_1, \ldots, x_n]$ tais que $f \cdot g \in \mathcal{I}(X)$ mas $f, g \notin \mathcal{I}(X)$. Dado $P \in X$, como $f(P)g(P) = 0$, ou $f(P) = 0$ ou $g(P) = 0$, de onde $X = (\mathcal{Z}(f) \cap X) \cup (\mathcal{Z}(g) \cap X)$. Mas se fosse o caso de $(\mathcal{Z}(f) \cap X) = X$, teríamos que $f \in \mathcal{I}(X)$, e idem para g. Para a recíproca, suponhamos que X é *redutível*, isto é, $X = A \cup B$ com A e B fechados distintos de X. Então existem $f \in \mathcal{I}(A)$ e $g \in \mathcal{I}(B)$ que testemunham essa diferença, isto é, tais que $f, g \notin \mathcal{I}(X)$. Por outro lado, dado $P \in X$, temos que $f(P)g(P) = 0$ pois um dos dois se anula em P, logo $fg \in \mathcal{I}(X)$. ■

Proposição 1.4.11. *Um espaço topológico noetheriano X pode ser escrito como uma união finita de subconjuntos fechados irredutíveis.*

Demonstração. A CCD nos permite fazer uma espécie de indução. Façamos em partes:
$*$ Vamos provar que se a afirmação vale para todo subconjunto fechado próprio de X, então vale para X. Se X for irredutível, não há o que se provar. Se X for redutível, então $X = Y \cup Z$ com $Y, Z \subsetneq X$, logo a união das decomposição para Y e Z dá uma decomposição para X.

$*$ Agora usamos a CCD: se o resultado não vale, existe um subconjunto fechado próprio de X onde também não vale, e assim sucessivamente, produzindo uma cadeia descendente de conjuntos fechados que não é estacionária. ■

Para conseguir unicidade nesta decomposição, basta retirar as "redundâncias", isto é, os conjuntos irredutíveis que estão inteiramente contidos em outros.:

Proposição 1.4.12. *Nas condições da proposição anterior, se* $X = Y_1 \cup \cdots \cup Y_r = Z_1 \cup \cdots \cup Z_s$, $Y_1, \ldots, Y_r, Z_1, \ldots, Z_r$ *são fechados irredutíveis de* X *tais que* $Y_i \subsetneq Y_j$ *e* $Z_i \subsetneq Z_j$ *se* $i \neq j$, *então* $\{Y_1, \ldots, Y_r\} = \{Z_1, \ldots, Z_r\}$.

Demonstração. Por simetria, basta mostrar que cada Z_i é algum Y_j. Como $Y_1 \cup \cdots \cup Y_r = Z_1 \cup \cdots \cup Z_s$, intersectando os dois lados por Z_i temos que $(Y_1 \cap Z_i) \cup \cdots \cup (Y_r \cap Z_i) = (Y_1 \cap Z_i)[(Y_2 \cap Z_i) \cup \cdots \cup (Y_r \cap Z_i)] = Z_i$. Como uniões e intersecções finitas de fechados são fechados e Z_i é irredutível, temos que $Z_i = Y_1 \cap Z_i$ ou $Z_i = (Y_2 \cap Z_i) \cup \cdots \cup (Y_r \cap Z_i)$. Assim existe j tal que $Z_i = Y_j \cap Z_i$, ou seja, $Z_i \subseteq Y_j$. Claro que o mesmo argumento mostra que existe i' com $Y_j \subseteq Z_{i'}$, o portanto temos $Z_i \subseteq Y_j \subseteq Z_{i'}$. Mas a propriedade de "retirar redundâncias" implica que $Z_i = Z_{i'}$, e portanto $Z_i = Y_j$, como queríamos demonstrar. ■

Analogia com os Ideais Radicais

O resultado a seguir vai ser usado mais tarde neste trabalho. Resolvemos colocar aqui pela semelhança com a proposição 1.4.11. Tal proposição pode ser vista no contexto de anéis da seguinte maneira: um ideal da forma $\mathcal{I}(X)$ pode ser escrito de maneira única como a intersecção de um número finito de ideais primos (da forma $\mathcal{I}(Y)$) não contidos um no outro. Na realidade, temos:

Proposição 1.4.13. *Sejam* A *um anel noetheriano e* $I \lhd A$ *radical. Então* I *é intersecção de um número finito de ideais primos.*

Demonstração. Novamente utilizaremos indução, apenas mudando o sentido da inclusão. Primeiro, provaremos que se I é um ideal radical tal que o resultado é válido para todo ideal radical J com $I \subsetneq J$, então o resultado também vale para I. A demonstração é concluída usando a CCA: a existência de um ideal radical que não satisfaz a conclusão desejada produz uma cadeia ascendente não estacionária desses ideais.

Seja I é um ideal radical tal que o resultado é válido para todo ideal radical acima dele. Se I é primo, nada a provar. Caso contrário existem $x, y \in A$ tais que $x, y \notin I$ mas $xy \in I$. Tomemos $J \doteq \sqrt{\langle I, x\rangle}$ e $K \doteq \sqrt{\langle I, y\rangle}$. Tanto J quanto K podem ser escritos como intersecção de um número finito de ideais primos. Basta mostrar que $I = J \cap K$. Que $I \subseteq J \cap K$ é imediato. Se $z \in J \cap K$, então, por estar em J existem $n \in \mathbb{N}$, $i_1 \in I$ e $a_1 \in A$ com $z^n = i_1 + a_1 x$, e por estar em K, existem $m \in \mathbb{N}$, $i_2 \in I$ e $a_2 \in A$ tais que $z^m = i_2 + a_2 y$. Logo $x^{n+m} = i_1 i_2 + i_1 a_2 y + i_2 a_1 x + a_1 a_2 xy \in I$. Como I é radical, $x \in I$. ■

Uma unicidade análoga ao do outro caso também pode ser demonstrada, mas não será necessária para os nossos propósitos.

Capítulo 2

Corpos com relações de ordem

O desenvolvimento da geometria algébrica se deu principalmente sobre corpos algebricamente fechados. No entanto, em $\mathbb{R}$ temos uma estrutura adicional para tratar o problema da falta de raízes: *a relação de ordem*.

Nós vamos lidar com uma generalização de $\mathbb{R}$, os chamados *corpos reais fechados*. A maior parte do será feito neste capítulo são fatos bem conhecidos sobre $\mathbb{R}$, o que pode fazer o leitor se perguntar se faz sentido ter esse trabalho todo para fazer uma generalização, já que provavelmente apenas $\mathbb{R}$ de fato vai ser usado nas aplicações. Para essa indagação temos duas respostas.

A primeira é que a teoria que apresentamos a seguir foi introduzida por Artin e Schreier para resolver o 17^o problema de Hilbert. Neste caso eles procuravam um corpo parecido com $\mathbb{R}$ mas que contivesse $\mathbb{R}(x)$.

A segunda resposta vai ser vista com mais detalhes no próximo capítulo. Essa generalização está escrita numa linguagem mais simples, em certo sentido, que a definição de $\mathbb{R}$, o que nos trás vantagens para provar alguns teoremas.

2.1 Geometria Semialgébrica

Vamos começar vendo um caso bastante conhecido:

Exemplo 2.1.1. O polinômio em 3 variáveis $p(X, A, B) \doteq X^2 + AX + B = 0$ possui raízes em $\mathbb{R}^3$ apenas na região em que $A^2 - 4B \geq 0$, e para cada par de valores para A e B satisfazendo esta inequação existe um valor para X tal que (X, A, B) satisfaz a equação original.

O exemplo acima nos mostra que para lidar com *igualdades* polinomiais em $\mathbb{R}$ somos convidados à aumentarmos um pouco o escopo de estudo e acrescentar as *desigualdades* polinomiais.

Além dos *conjuntos algébricos* definidos no capítulo anterior, aqui são naturais os conjuntos *semi-algébricos*:

Definição 2.1.2. Seja k um corpo ordenado. Um subconjunto X de k^n é dito semialgébrico se ele pode ser construído por uniões e intersecções finitas de conjuntos da forma $\{p \in k^n \mid f(p) = 0\}$ ou da forma $\{p \in k^n \mid f(p) > 0\}$, onde $f(x_1, \ldots, x_n) \in k[x_1, \ldots, x_n]$.

O exemplo 2.1.1 é uma instância de uma das formulações de resultado muito importante:

Teorema 2.1.3 (Tarski-Seidenberg geométrico)**.** *A projeção de um conjunto semialgébrico em $\mathbb{R}^n$ sobre k quaisquer coordenadas é um conjunto semialgébrico de $\mathbb{R}^k$.*

Esse resultado é consequência imediata do teorema de eliminação de quantificadores para a teoria dos corpos reais fechados (4.2.9). No momento, iremos abordar um pouco mais a teoria dos

corpos em que vamos trabalhar: corpos *ordenados*, *ordenáveis* e, principalmente, os *corpos reais fechados*.

2.2 Corpos formalmente reais

Quando falamos em corpos ordenados, esperamos alguma compatibilidade da relação de ordem com a estrutura de corpo.

Definição 2.2.1. Dizemos que um corpo R é **ordenado** se além das operações de soma e produto ele está munido de uma relação de ordem total $\leq$ que satisfaz os seguintes axiomas:

i) $\forall x, y, z \in R$, $x = y \implies x + z \leq y + z$

ii) $\forall x, y, z \in R$, $0 \leq z$ e $x \leq y \implies xz \leq yz$

Por abuso de linguagem, vamos chamar de *ordem* em um corpo apenas as relações de ordem que o tornam um *corpo ordenado*, e chamaremos de *corpo formalmente real* ou *corpo ordenável* a um corpo que admitir *alguma* ordem.

Definição 2.2.2. Seja $\langle k, \leq \rangle$ um corpo ordenado.
a) Para $a \in k$, definimos (como é usual)

$$|a| = \begin{cases} a & \text{se } a \geq 0; \\ -a & \text{se } a < 0, \end{cases}$$

denominado **módulo** de a em $\langle k, \leq \rangle$. Note que, para todo $b \in k$, $|b| \geq 0$.

b) Se $a < b$ em k, $(a, b) = \{c \in k : a < c < b\}$ é o **intervalo aberto** de extremidades a, b. Analogamente, definimos intervalo fechado, $[a, b]$.

Lema 2.2.3. *Seja $\langle k, \leq \rangle$ um corpo ordenado.*
a) Para a, $b \in k$, temos

(1) $|a| = 0 \Leftrightarrow a = 0$; (2) $|ab| = |a|\ |b|$; (3) $|a + b| \leq |a| + |b|$;

(4) $\big||a| - |b|\big| \leq |a - b|$.

b) Sejam $p(x) \in k[x]$ um polinômio com coeficientes em k, $r > 0$ e $I = [-r, r]$. Então:

(1) *$p(x)$ é limitado em I, i.e., existe $M > 0$ tal que para todo $b \in I$, $|p(b)| \leq M$;*

(2) *A restrição de $p(x)$ ao intervalo I é uniformemente contínua na topologia da ordem em k, i.e., para todo $\varepsilon > 0$, existe $\delta > 0$, tal que para a, $b \in I$, $|a - b| < \delta \ \Rightarrow\ |p(a) - p(b)| < \varepsilon$.*

(3) (Preservação do sinal) *Se $p(a) \neq 0$, então existe $\delta > 0$ tal que para todo $b \in (a - \delta, a + \delta)$, temos $p(a)p(b) > 0$.*

Demonstração. a) A mesma prova utilizada nos cursos elementares funciona neste caso.

b) Seja $p(x) = \sum_{k=0}^{n} a_k x^k$.

(1) Para $b \in I$, a desigualdade triangular ((a).(3)) fornece

$$|p(b)| \leq \textstyle\sum_{k=0}^{n} |a_k|\ |b|^k \leq \sum_{k=0}^{n} |a_k|\ r^k,$$

como desejado.

(2) Para a, $b \in I$, Note que $p(b) - p(a)$ é divisível por $b - a$, i.e., existe um polinômio $q(x)$ tal que $p(b) - p(a) = (b - a)(q(b) - q(a))$. Por (b).(1), existe $C > 0$ tal que $|q(b) - q(a)| \leq C$ e portanto

$$|p(b) - p(a)| \leq |b - a|\ C,$$

e a conclusão desejada segue imediatamente desta desigualdade, tomando $\delta = \varepsilon/(1 + C)$.

O item (3) segue de (2) com argumento idêntico ao utilizado em cursos elementares de Cálculo. ■

2.2.1 Cones

Uma ordem em um corpo está intrinsecamente ligada ao conjunto dos elementos positivos. Melhor dizendo, apenas sabendo o subconjunto P (de um corpo k) dos elementos positivos podemos recuperar a relação de ordem pois $x \leq y$ se, e somente se, $y - x \in P$.

Notação 2.2.4. Dada uma ordem ordem de $\leq$ em um corpo R, denotaremos o conjunto de seus elementos positivos $\{x \in R | x \geq 0\}$ por $P_{\leq}$.

Agora vamos extrair algumas propriedades essenciais de um "conjunto dos elementos positivos".

Definição 2.2.5. Um subconjunto C de um corpo k é dito um *cone* se

(i) $x, y \in C \implies x + y \in C$ e $xy \in C$

(ii) $x \in k \implies x^2 \in C$

Além disso, se $-1 \notin C$, C é dito um cone *próprio*.

Em corpos de característica diferente de 2 (como, veremos, é sempre o caso dos corpos ordenáveis), essa definição de *próprio* coincide com a definição conjuntista: se $-1 \in C$ então, qualquer que seja $x \in k$, $x = [(1 + x)^2 - (1 - x)^2] \cdot (\frac{1}{2})^2 \in C$.

Notação 2.2.6. Todo cone contém o cone $\sum k^2 \doteq \{x \in k | \exists a_1, \cdots, a_n \in k \text{ com } x = a_1^2 + \cdots + a_n^2\}$.

A verificação de que de fato $\sum k^2$ é um cone é trivial.

Lema 2.2.7. *Seja C um cone próprio de um corpo k. Então:*

(i) Se $a \in C$ e $a \neq 0$, então $\frac{1}{a} \in C$.

(ii) $C \cap -C = 0$.

Demonstração. Para (i), temos que $\frac{1}{a} = a \cdot (\frac{1}{a})^2 \in C$; para (ii), se existesse $x \neq 0$ com $x \in C \cap -C$, teríamos $-1 = (-x) \cdot \frac{1}{x} \in C$. ■

Proposição 2.2.8. *Seja $(R, \leq)$ um corpo ordenado. Então $P_{\leq}$ é um cone próprio tal que $R = P_{\leq} \cup -P_{\leq}$. Reciprocamente, se C é um cone próprio de um corpo k que satisfaz $C \cup -C$, então existe uma ordem $\leq$ em k, com $C = P_{\leq}$.*

Demonstração. Se $x \leq 0$, somando $-x$ aos dois lados da desigualdade, temos que $0 \leq -x$. Assim $R = P_{\leq} \cup -P_{\leq}$. Para provarmos que $P_{\leq}$ é um cone, a única propriedade que não é direta é de conter todo os quadrados. Se $x \in P_{\leq}$, $x \cdot x \in P_{\leq}$. Caso contrário, $-x \in P_{\leq}$ e caímos no caso anterior, pois $(-x)^2 = x^2$. Para mostrar que $P_{\leq}$ é próprio, basta notar que, então, $1 = 1^2 > 0$, logo $0 > -1$.

Para a recíproca, definamos $\leq$ por $x \leq y \iff y - x \in C$. Verifiquemos as propriedades de ordem:

(i) *(reflexiva)* Como $0 = 0^2 \in C$, para todo $x \in k$, $x \leq x$;

(ii) *(antissimétrica)* Pelo lema 2.2.7, $C \cap -C = 0$. Portanto se $x - y \in C$ e $y - x \in C$, $x = y$.

(iii) *(transitiva)* Basta somar: $(x - y) \in C$ e $(y - x) \in C \implies (x - z) \in C$.

Além disso a ordem é total pois, se $a, b \in k$, ou $a - b \in C$ ou $b - a \in C$. A compatibilidade da ordem com as operações segue diretamente da definição de cone. ■

O leitor deve estar acostumado com a ordem em $\mathbb{R}$ e seus subcorpos. Com o próximo resultado vamos ter exemplos de ordens que não são arquimedianas.

Proposição 2.2.9. *Seja $(R, \leq)$ um corpo ordenado. Existe uma única ordem em $R(x)$ que estende a ordem de R e tal que x é positivo mas menor do que todo elemento estritamente positivo de R.*

Demonstração. **Unicidade, supondo que existe uma ordem com essas propriedades:** Primeiramente, é fácil ver que $x^n < \cdots < x^2 < x$, sendo que todos são positivos. Agora vejamos que $bx^n < a$ para todos $a, b \in R$ com $a > 0$. Se $b \leq 0$, não há o que fazer. Se $b > 0$, basta usar que $x^n < x < \frac{a}{b}$. Com isso conseguimos que todo polinômio da forma $x \cdot p(x)$, com $p(x) \in k[x]$, é menor que todo elemento estritamente positivo a de k (basta notar que cada monômio é menor que $\frac{a}{n+1}$, onde n é o grau de $p(x)$).

Tomemos agora um polinômio $f(x) = x \cdot p(x) + c$ com termo constante c não nulo. Afirmo que o sinal de $f(x)$ é sinal de c. Comecemos supondo $c > 0$. Se $f(x)$ fosse negativo teríamos que $x \cdot p(x) < -c$, de onde $x \cdot [-p(x)] > c$, um absurdo. Se c fosse negativo, então $-f(x)$ seria positivo e portanto $f(x)$, negativo. Se o termo constante de um polinômio for zero, basta olhar para o coeficiente do termo de menor grau (digamos, k), pois temos que o sinal de $x^k \cdot p(x)$ é bem determinado.

Até agora vimos que o sinal dos polinômios estão todos determinados. Falta ver o das funções racionais. Quando temos que $\frac{p(x)}{q(x)} > 0$? Se $q(x) > 0$, se e somente se $p(x) > 0$ e se $q(x) < 0$, se e somente se $q(x) < 0$. Assim o sinal de $\frac{p(x)}{q(x)}$ é sempre o mesmo de $p(x) \cdot q(x)$.

Existência: Nossa única esperança é definir o sinal um polinômio $p(x) = a_n x^n + \cdots + a_k x^k$ como sendo o mesmo de $a_k (\neq 0)$, e de uma função racional $\frac{p(x)}{q(x)}$ como sendo o sinal de $p(x) \cdot q(x)$ (note que este sinal não depende da fração escolhida).

Chamemos o conjunto dos elementos positivos de C. Se uma fração $\frac{p(x)}{q(x)}$ não for positiva então $-\frac{p(x)}{q(x)}$ será, de onde $C \cup -C = R(x)$. Falta só ver que C é um cone:

É claro que a soma e o produto de dois polinômios positivos dá positivo. Também é claro que o quadrado de um polinômio é positivo. Para ver as funções racionais, basta escolher frações com denominadores positivos, pois temos:

$$\frac{p(x)}{q(x)} + \frac{f(x)}{g(x)} = \frac{p(x)q(x)+f(x)q(x)}{q(x)g(x)} \quad \text{e} \quad \frac{p(x)}{q(x)} \cdot \frac{f(x)}{g(x)} = \frac{p(x)f(x)}{q(x)g(x)}$$

■

Lema 2.2.10. *Sejam k um corpo e C um cone próprio de k. Então*

(i) Se $-a \notin C$, então $C[a] = \{x + ay | x, y \in C\}$ é um cone próprio de k.

(ii) O cone C está contido no conjunto dos pontos positivos de alguma ordem de k (em particular, se existe algum cone próprio em k, k é ordenável).

Demonstração. (i) Que $C[a]$ é fechado pela soma e que todos os quadrados de k estão em $C[a]$ é trivial. Quanto ao produto, se $x, x', y, y' \in C$ temos $(x+ay)(x'+ay') = (xx'+a^2yy')+a(xy'+x'y) \in C[a]$. Se $-1 = x + ay$, com $x, y \in C$, então ou $y = 0$ e $-1 \in C$ ou $y \neq 0$ e $-a = \frac{x+1}{y} = x\frac{1}{y} + \frac{1}{y} \in C$.

(ii) O Lema de Zorn nos garante que existe um cone próprio maximal P contendo C. Com efeito, seja $\mathfrak{C}$ uma cadeia de cones próprios contendo C. Certamente $\bigcup \mathfrak{C}$ é fechado por soma e produto, $\sum k^2 \subseteq \bigcup \mathfrak{C}$ e $-1 \notin \bigcup \mathfrak{C}$.

Seja $a \in k$ com $a \notin P$. Como $P[-a]$ é um cone próprio contendo P e P é maximal, $P[-a] = P$ e assim $-a \in P$, *i.e.*, $a \in -P$. A proposição 2.2.8 conclui a demonstração. ■

Teorema 2.2.11. *Seja k um corpo. São equivalentes:*

(i) k é ordenável; *(ii) k possui um cone próprio.* *(iii) $-1 \notin \sum k^2$;*

(iv)Se $x_1, \cdots, x_n \in k$ são tais que $\sum_{i=1}^n x_i^2 = 0$, então $x_1 = \cdots = x_n = 0$.

Demonstração. Já temos que $i) \implies ii) \implies iii) \implies i)$. Só falta $iii) \iff iv)$. Ora, se $-1 = a_1^2 + \cdots + a_r^2 \in \sum k^2$, então $0 = 1^2 + a_1^2 + \cdots + a_r^2$, mas $1 \neq 0$. Reciprocamente, se existe $\sum_{i=1}^n x_i^2 = 0$ com $x_1 \neq 0$, então $-1 = \sum_{i=2}^n (\frac{x_i}{x_1})^2$. ■

Vale notar que, em particular, todo corpo ordenável tem caraterística 0, pois para todo p primo, $\sum_{i=1}^p 1 = 0$ implica $-1 = \sum_{i=1}^{p-1} 1^2$.

Proposição 2.2.12. *Sejam k um corpo ordenável, C um cone próprio de k e $\{\leq_\lambda\}_{\lambda\in\Lambda}$ a família de todas as ordens de k tais que $C \subseteq P_{\leq_\lambda}$. Então $C = \bigcap_{\lambda\in\Lambda} P_{\leq_\lambda}$.*

Demonstração. É claro que $C \subseteq \bigcap_{\lambda\in\Lambda} P_{\leq_\lambda}$. Se $a \notin C$, existe (por 2.2.10) uma ordem $\leq_{\lambda_0}$ tal que $C[-a] \subseteq P_{\leq_{\lambda_0}}$. Como $a \neq 0$ e $-a \in P_{\leq_{\lambda_0}}$, $a \notin P_{\leq_{\lambda_0}}$ e, portanto, $a \notin \bigcap_{\lambda\in\Lambda} P_{\leq_\lambda}$. ■

2.3 Corpos Reais Fechados

Corpos reais fechados estão para a Geometria Algébrica Real assim como corpos algebricamente fechados estão para a Geometria Algébrica clássica. Eles são os corpos que mantém todas as características "algébricas" de $\mathbb{R}$ (mais precisamente, a Teoria de 1^a ordem na linguagem dos anéis ordenados, como veremos no capítulo 4) e são, entre os corpos ordenáveis, os que mais se aproximam de um corpo algebricamente fechado.

Definição 2.3.1. Um corpo formalmente real R é dito **real fechado** se não possui nenhuma extensão algébrica própria que também seja formalmente real.

A definição foi escolhida entre muitas afirmações equivalentes, como veremos a seguir. Mas antes, vejamos algumas propriedades destes corpos. No restante da seção, R indica um corpo real fechado fixado.

Proposição 2.3.2. *Existe uma única ordem que torna R um corpo ordenado e, nesta ordem, todo elemento positivo é um quadrado.*

Demonstração. Seja $a \in R$. Se a não é um quadrado perfeito, então a extensão $R[\sqrt{a}]$ é algébrica e própria, logo não é formalmente real. Em particular podemos escrever -1 como soma de quadrados em $R[\sqrt{a}]$:

$$-1 = \sum_{k=1}^n (x_k + y_k\sqrt{a})^2 = \sum_{k=1}^n x_k^2 + a \cdot \sum_{k=1}^n y_k^2.$$

Se a soma $\sum_{k=1}^n y_k^2$ fosse nula, teríamos que R não é formalmente real. Logo podemos escrever

$$-a = \left(\sum_{k=1}^n y_k^2\right)^{-1} \cdot \left(1 + \sum_{k=1}^n x_k^2\right) \in \sum R^2.$$

Desta equação decorrem dois fatos:

- O cone $\sum R^2$ é tal que $R = \sum R^2 \cup -\sum R^2$, logo define uma ordem em R, a qual tem que ser única.
- Se a não é quadrado, ele é um elemento negativo nessa ordem. ■

Corolário 2.3.3. *Todo corpo real fechado é um corpo* pitagórico*,* i.e.*, toda soma de quadrados é um quadrado.*

Proposição 2.3.4. *Todo polinômio de grau ímpar em $R[x]$ possuí alguma raiz em R.*

Demonstração. Suponha que existam polinômios de grau ímpar sem raiz em R e seja $f \in R[x]$ um desses polinômios, mas cujo o grau é *mínimo* entre esses.

Como o grau de f é ímpar, decompondo f em polinômios irredutíveis encontramos um fator de grau ímpar $d > 1$ também sem raízes, logo, pela minimalidade do grau, f é irredutível. Assim $\frac{R[x]}{\langle f\rangle}$ é uma extensão algébrica não trivial de R, de onde concluímos que existem $\bar{h}_1, \ldots, \bar{h}_m \in \frac{R[x]}{\langle f\rangle}$ tais que

$$-1 = \sum_{k=1}^{m} \bar{h}_k^2 \in \tfrac{R[x]}{\langle f\rangle}$$

o que implica

$$-1 = \sum_{k=1}^{m} h_k^2 + f \cdot g \in R[x],$$

para certos $g, h_1, \ldots, h_m \in R[x]$, com os graus de h_k sempre menores que d.

Mas então o grau de $\sum_{k=1}^{m} h_k^2$ é menor ou igual a $2d-2$, e certamente é par, uma vez que o coeficiente do termo de maior grau é soma dos de alguns dos coeficientes dos termos de maior grau dos h_k^2 (esta soma não se anula pois é soma de quadrados em um corpo formalmente real).

Sendo assim, o grau de g tem que ser ímpar e menor ou igual a $d-2$. Mas então g possui uma raiz $r \in R$, de onde

$$-1 = \sum_{k=1}^{m} h_k(k)^2 \in R,$$

contradizendo o fato de que R é formalmente real. ■

Na próxima seção vamos provar que se um corpo ordenado é tal que todo elemento positivo tem raiz quadrada e que todo polinômio de grau ímpar possui raiz no próprio corpo, então vale o *Teorema Fundamental da Álgebra*, ou, melhor dizendo, que $R[\sqrt{-1}] = \frac{R[x]}{\langle x^2+1\rangle}$ é um corpo algebricamente fechado[1] (teorema 2.4.9). Mas vamos assumir este fato por enquanto para tirarmos mais alguns resultados. Também vamos fazer a convenção de denotar por i uma raiz quadrada fixada de -1.

Uma coisa que vamos usar é que se $a+bi$ é raiz de um polinômio com coeficientes em R, então $a-bi$ também o é. Isto é uma aplicação direta do seguinte fato, cuja prova é uma simples verificação:

Proposição 2.3.5. *Seja k um corpo onde -1 não tem raiz quadrada; então a função $\sigma : k[i] \longrightarrow k[i]$, dada por $\sigma(a+bi) = a-bi$, é um homomorfismo de corpos.*

Proposição 2.3.6 (Teorema do Valor Intermediário para Polinômios)**.** *Seja $p(x) \in R[x]$ um polinômio tal que existem $a, b \in R$ com $p(a) < 0$ e $p(b) > 0$. Então $p(x)$ possui uma raiz em R, situada entre a e b.*

[1] Note que as provas usuais usando análise e/ou topologia para provar que $\mathbb{C}$ é algebricamente fechado não podem ser usadas diretamente neste contexto

Demonstração. Como $R[i]$ é algebricamente fechado, os polinômios irredutíveis em $R[x]$ são ou lineares ou são da $(x-c-di)(x-c+di) = (x-c)^2+d^2$, os quais nunca mudam de sinal. Assim, se $p(x)$ troca de sinal da a para b é porque um de seus fatores irredutíveis *lineares* o faz. Mas para polinômios de grau 1 o resultado é trivial. ■

Com esse fato, temos outro modo para mostrar que todo elemento positivo tem raiz quadrada e que todo polinômio de grau ímpar tem raiz em R. Vamos usar isso para obter algumas equivalências.

Lema 2.3.7. *Seja $(k,\leq)$ um corpo ordenado e $f(x) = a_nx^n + \cdots + a_0 \in k[x]$, com $a_n \neq 0$. Então existe $M \in k$, $M > 0$, tal que se $|c| \geq M$, o sinal de $f(c)$ é o mesmo sinal de a_nc^n.*

Demonstração. Podemos escrever $f(x)$ como $a_nx^n(1 + b_1x^{-1} + \cdots + b_nx^{-n})$, com $b_1, \ldots, b_n \in k$. Defino $M \doteq 1 + |b_1| + \cdots + |b_n| \geq 1$. Então, se $|c| \geq M$:

$$|b_1c^{-1} + \cdots + b_nc^{-n}| \leq |b_1c^{-1}| + \cdots + |c_nx^{-n}| \leq \ (|b_1| + \cdots + |b_n|)|c|^{-1} < 1.$$

Logo, o valor de $|a_nc^n|$ é maior do que o de $|f(c) - ac^n|$ e assim o sinal depende apenas do termo de grau máximo. ■

Teorema 2.3.8. *Seja k um corpo. São equivalentes:*

(i) k é real fechado;

(ii) Existe uma única ordem que torna k um corpo ordenado, e nesta ordem todo elemento positivo é um quadrado. Além disso, todo polinômio em $k[x]$ de grau ímpar possui raiz em k;

(iii) O anel quociente $\frac{k[x]}{(x^2+1)}$ é um corpo algebricamente fechado;

(iv) Existe uma única ordem $\leq$ que torna k um corpo ordenado, e uma condição suficiente para um polinômio $p(x)$ ter raíz em k é existirem $a, b \in R$ com $p(a) < 0$ e $p(b) > 0$.

Demonstração. Já provamos $(i) \implies (ii)$ e vamos deixar $(ii) \implies (iii)$ para a próxima seção. Assumindo este último, também temos que $(i) \implies (iv)$. Para fechar as equivalências, basta provar as implicações que seguem:

• $(iii) \implies (i)$: Como $\frac{k[x]}{(x^2+1)}$ é corpo, -1 não é um quadrado em k. E por ser uma extensão algebricamente fechada de grau primo sobre k, é a única extensão algébrica de k. Só resta então mostrar que k é formalmente real, isto é, que -1 não é soma de quadrados. Já sabemos que não é um quadrado. Vamos concluir mostrando que toda soma de quadrados é um quadrado. Sejam $a, b \in k$. Temos que $a+bi$ tem raiz quadrada em $k[i]$, *i.e.*, existem $c, d \in k$ tais que $a+bi = (c+di)^2$. Multiplicando pelo conjugado, $a^2 + b^2 = (c^2 + d^2)^2$.

• $(iv) \implies (iii)$: Se $a \in k$ é um elemento positivo, como polinômio $x^2 - a$ tem sinais diferentes em 0 e $a+1$, a tem raiz quadrada. O lema 2.3.7 garante que todo polinômio de grau ímpar troca de sinal, logo tem raiz em k. ■

2.4 O Teorema Fundamental da Álgebra

Iremos, basicamente, seguir a prova dada por Laplace em 1795, antes mesmo da primeira prova dada por Gauss (1799). Na época Gauss fez críticas pertinentes às "provas" deste teorema que já existiam, mostrando seus erros ou suas hipóteses escondidas. A de Laplace (bem como a do próprio Gauss) estava incompleta. Laplace assumia que existiam raízes do polinômio *em algum lugar* e

depois provava que estas eram complexas. Hoje em dia o problema está resolvido, usando a bem conhecida construção do *corpo de raízes de um polinômio*.

O principal pré-requisito para a prova de Laplace é o Teorema Fundamental dos Polinômios Simétricos, que será o tema da subseção a seguir.

2.4.1 Polinômios Simétricos

Sejam k um corpo e $a_1, \cdots, a_n \in k$. Então, os coeficientes do polinômio

$$p(x) \doteq \prod_{1\leq i\leq n}(x - a_i) = x^n + \sum_{1\leq j\leq n}(-1)^j \lambda_j x^{n-j}$$

são tais que

$$\lambda_j = \sum_{\substack{J\subseteq\{1,\cdots,n\}\\ |J|=j}} \prod_{i\in J} a_i.$$

Essas relações entre coeficientes e raízes, conhecidas como *Relações de Girrard*, nos servirão para passar certas funções nas raízes de $p(x)$ para funções nos coeficientes de $p(x)$ de maneira bastante agradável.

Definição 2.4.1. Um polinômio $Q(X_1, \cdots, X_n) \in k[X_1, \cdots, X_n]$ é dito **simétrico** se para toda permutação $\sigma : \{1, \cdots, n\} \to \{1, \cdots, n\}$, $\ Q(X_1, \cdots, X_n) = Q(X_{\sigma(1)}, \cdots, X_{\sigma(n)})$.

Definição 2.4.2. Chamaremos os polinômios

$$E_j \doteq \sum_{\substack{J\subseteq\{1,\cdots,n\}\\ |J|=j}} \prod_{i\in J} X_i \in k[X_1, \cdots, X_n],$$

com $j \in \mathbb{N}$, $1 \leq j \leq n$, de **polinômios simétricos elementares**.

Claramente, todos os polinômios simétricos elementares são, de fato, simétricos. O resultado que utilizaremos é o seguinte:

Teorema 2.4.3 (Fundamental dos Polinômios Simétricos)**.** *Se $Q \in k[X_1, \cdots, X_n]$ é um polinômio simétrico então existe $f \in k[X_1, \cdots, X_n]$, não necessariamente simétrico, tal que*

$$Q = f(E_1, \cdots, E_n).$$

Mais precisamente, ele vai ser útil da seguinte forma:

Corolário 2.4.4. *Seja $p \in k[x]$ um polinômio de grau n, e sejam $a_1, \ldots, a_n$ as raízes de p (com multiplicidade) em algum corpo de raízes $F \supseteq k$. Se $Q \in k[X_1, \cdots, X_n]$ é um polinômio simétrico, então $Q(a_1, \cdots, a_n) \in k$.*

O ponto fundamental da prova do Teorema 2.4.3 consiste em ter uma maneira razoável de organizar os monômios de um polinômio simétrico.

Se um polinômio simétrico possui um monômio $c \cdot X_1^{\alpha_1} \cdots X_n^{\alpha_n}$, ele necessariamente tem os monômios $c \cdot X_1^{\alpha_{\sigma(1)}} \cdots X_n^{\alpha_{\sigma(n)}}$, para todo $\sigma \in S_n$. Assim, dado $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}$ com $\alpha_1 \geq \cdots \geq \alpha_n$ podemos definir $X^\alpha \doteq \sum_{\sigma\in S_n} X_1^{\alpha_{\sigma(1)}} \cdots X_n^{\alpha_{\sigma(n)}}$ e teremos que um polinômio simétrico será uma soma de elementos da forma $c_\alpha \cdot X^\alpha$.

Notação 2.4.5. Denotaremos o conjunto $\{(\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n \mid \alpha_1 \geq \cdots \geq \alpha_n\}$ por $\mathcal{N}$.

Agora vamos colocar a ordem *lexicográfica* nas possíveis n-uplas: se $\alpha = (\alpha_1, \ldots, \alpha_n), \beta = (\beta_1, \ldots, \beta_n) \in \mathbb{N}^n$, diremos que α é *maior* que β se existe $k \in \{1, \ldots, n\}$ tal que $\alpha_i = \beta_i$ se $i < k$ e $\alpha_k > \beta_k$.

Lema 2.4.6. *Com a ordem definida acima, se* $\alpha^1 \geq \alpha^2 \geq \cdots \geq \alpha^k \geq \cdots$ *é uma sequência de elementos*[2] *de* $\mathcal{N}$*, então existe* $N \in \mathbb{N}$ *tal que* $\alpha^k = \alpha^N$*, para todo* $k \geq N$

Demonstração. Basta notar que o conjunto dos elementos menores que $\alpha^1 = (\alpha_1^1, \ldots, \alpha_n^1)$ tem no máximo $(\alpha_1^1)^n$ elementos. ■

Demonstração do Teorema 2.4.3. Seja $Q(X_1, \ldots, X_n)$ um polinômio simétrico em $k[X_1, \ldots, X_n]$ cujo "termo" de maior índice é $c \cdot X^\alpha \neq 0$. Este também é o "termo" com o maior índice de $c \cdot \prod_{i=1}^n E_i^{\alpha_i - \alpha_{i+1}}$ (pondo $\alpha_{n+1} \doteq 0$). Subtraindo o segundo do primeiro, o problema se reduz a um polinômio cujo "maior termo" tem índice estritamente menor do que α.

Basta aplicarmos este procedimento um número finito de vezes para chegarmos ao polinômio nulo, pois caso contrário, conseguiríamos uma sequência infinita de elementos de $\mathcal{N}$, estritamente decrescente. ■

2.4.2 Prova do Teorema Fundamental da Álgebra

No que se segue vamos fixar R como sendo um corpo ordenado onde todo elemento positivo é um quadrado e todo polinômio em $R[x]$ com grau ímpar tem raiz em R.

Vamos começar provando o teorema para polinômios de grau 2 e com uma "continha" muito conhecida.

Proposição 2.4.7. *Seja* k *um corpo com característica diferente de* 2*. O polinômio* $p(x) = x^2 + ax + b \in k[x]$ *pode ser escrito como* $(x + c)^2 + d$*, com* $c, d \in k$ *e tem raízes em* k *se, e somente, se* $\Delta \doteq a^2 - 4b$ *é um quadrado em* k*.*

Demonstração. Completando quadrados, obtemos:

$$\begin{aligned} x^2 + ax + b &= x^2 + ax + \left(\tfrac{a}{2}\right)^2 - \left(\tfrac{a}{2}\right)^2 + b = \left(x + \tfrac{a}{2}\right)^2 - \left(\tfrac{a}{2}\right)^2 + b = \left(x + \tfrac{a}{2}\right)^2 + \tfrac{-a^2+4b}{4} \\ &= \left(x + \tfrac{a}{2}\right)^2 - \tfrac{\Delta}{4}. \end{aligned}$$

■

Corolário 2.4.8. *Todo polinômio de grau dois com coeficientes em* $R[i]$ *tem raiz e todas as suas raízes estão em* $R[i]$*.*

Demonstração. Basta verificar que todo elemento de $R[i]$ é um quadrado. Seja $a + bi \in R[i]$, com $a, b \in R$. Vamos mostrar que uma de suas raízes quadradas é $\sqrt{\frac{a+\sqrt{a^2+b^2}}{2}} + sgn(b)\sqrt{\frac{-a+\sqrt{a^2+b^2}}{2}}i$ (onde $sgn(b) = 1$ se $b \geq 0$ e $sgn(b) = -1$ se $b < 0$).

Note que todos os números dentro de radicais acima são positivos e suas raízes estão em R. Um cálculo simples, mostrará que

$$\left(\sqrt{\tfrac{a+\sqrt{a^2+b^2}}{2}} + sgn(b)\sqrt{\tfrac{-a+\sqrt{a^2+b^2}}{2}}i\right)^2 = \tfrac{2a}{a} + sgn(b)\sqrt{b^2}i = a + bi,$$

como necessário. ■

Tendo o caso de grau 2 em mãos, a estratégia para o caso geral é uma indução em "quão par é o grau" do polinômio. Melhor explicando, se o grau do polinômio é $2^n m$, com m ímpar, vamos fazer uma indução em n, por mais que o próprio grau do polinômio aumente.

Teorema 2.4.9 (Fundamental da Álgebra)**.** *O corpo* $R[i]$ *é algebricamente fechado.*

[2] O resultado ainda seria válido se colocássemos elementos de $\mathbb{N}^n$, mas isto é desnecessário para o nosso caso

Demonstração. Vamos começar com $f \in R[x]$, cujo grau é $2^n m$ com m ímpar. Vamos provar que f tem raiz em $R[i]$ por indução em n. Se $n = 0$, f tem raiz em R. Suponhamos que $n > 0$.

Sejam $a_1, \ldots, a_{2^n m}$ as raízes de f, contadas com multiplicidades, em algum corpo de raízes de f. Vamos definir uma sequência (infinita) de polinômios: para cada $k \in \mathbb{N}$, seja

$$p_k(x) = \prod_{1 \leq i < j \leq 2^n m} [x - (a_i + a_j + k a_i a_j)].$$

Os coeficientes de cada p_k não dependem da ordenação dada para as raízes e são polinomiais nelas, logo, pelo corolário 2.4.4 do Teorema Fundamental dos Polinômios Simétricos, os coeficientes de p_k estão em R.

O grau de cada p_k é $\binom{2^n m}{2} = 2^{n-1} k(2^n k - 1)$. Como $k(2^n k - 1)$ é ímpar, pela hipótese de indução, cada um dos p_k tem pelo menos uma raiz em $R[i]$. Uma vez que há infinitos polinômios da forma p_k, existe pelo menos um par a_i, a_j de raízes de f tal que $(a_i + a_j + k a_i a_j)$ é raiz em $R[i]$ de p_k para 2 valores distintos de k. Assim, tanto $a_i + a_j$ quanto $a_i a_j$ estão em $R[i]$. Mas a_i e a_j são as raízes de $x^2 - (a_i + a_j)x + a_i a_j \in R[i][x]$. Por 2.4.8, todas as raízes deste polinômio estão em $R[i]$ e, portanto, a_i e a_j estão em $R[i]$.

Por enquanto, estabelecemos que um polinômio com coeficientes em R tem raiz em $R[i]$. Para um polinômio com coeficientes em $R[i]$, usamos a "*conjugação complexa*" (2.3.5). Para cada $g \in R[i][x]$, podemos escrever $g = p + iq$, com $p, q \in R[x]$. Definimos $\bar{g} \doteq p - iq$. Assim, o polinômio $g\bar{g} = p^2 + q^2 \in R[x]$ e, portanto, possui uma raiz $\alpha \in R[i]$. Se α é raiz de g, acabou. Se α é raiz de $\bar{g}$, então, como conjugação é homomorfismo, $\bar{\alpha}$ é raiz de g. ■

2.4.3 Recíprocas para o Teorema Fundamental da Álgebra

Já vimos que se R é um corpo tal que $\frac{R[x]}{(x^2+1)}$ é um corpo algebricamente fechado, então R é real fechado (2.3.8). Ainda temos outras "recíprocas" para o Teorema Fundamental da Álgebra.

Teorema 2.4.10. *Se k é um corpo algebricamente fechado de característica zero, então existe um subcorpo $R \subseteq k$ o qual é real fechado e tal que $k = R[i]$.*

Demonstração. Como a característica de k é zero, existe uma cópia de $\mathbb{Q}$ contida em k, a qual é formalmente real. Seja $\mathcal{C}$ o conjunto dos subcorpos formalmente de k, parcialmente ordenado pela inclusão. É fácil verificar que com esta ordem parcial $\mathcal{C}$ satisfaz as hipóteses do Lema de Zorn. Se R é um elemento maximal em $\mathcal{C}$, então, uma vez que toda extensão algébrica pode imersa em k, R é um corpo formalmente real que não possui extensão algébrica não trivial formalmente real, mostrando que R é real fechado e que $R[i]$ é algebricamente fechado (por 2.4.9).

Para terminar o argumento, basta mostrar que k é algébrico sobre R, e portanto seu fecho algébrico. Seja $a \in k$. Se a não for algébrico sobre R, então o corpo $R(a)$ é isomorfo a um corpo de funções racionais sobre um corpo ordenado, e portanto ordenável (ver 2.2.9), o que contradiz a maximalidade de R em $\mathcal{C}$. ■

Um resultado surpreendente é que se k é um corpo *não* algebricamente fechado (*a priori* de qualquer característica) mas cujo fecho algébrico é uma extensão *finita*, então k é real fechado. Não vamos provar (nem usar) tal teorema neste texto, mas o leitor interessado poderá encontrar a prova em [5] e [6].

Capítulo 3

Cálculo Diferencial e o Teorema de Sturm

Neste capítulo, mostraremos que muitos dos resultados do Cálculo Diferencial para polinômios em uma variável em $\mathbb{R}$ continuam valendo sobre um corpo real fechado qualquer, usando o conceito de *derivada formal*.

Terminaremos com resultados sobre contagem de raízes e fecho real, em particular um importante resultado devido a Sturm.

O leitor provavelmente já se deparou com o conceito de derivada *formal* de um polinômio. Neste capítulo, mostraremos que muitos dos resultados do Cálculo Diferencial podem ser transportados para polinômios em uma variável sobre um corpo real fechado, usando a derivada formal. No decorrer desse capítulo, k denotará um corpo qualquer e R um corpo real fechado.

3.1 Derivada Formal

Vamos definir derivada formal tentando imitar a definição de derivada na Análise:

Definição 3.1.1 (Derivada)**.** Sejam k um corpo e $f(x) \in k[x]$ um polinômio e y uma variável algebricamente independente sobre $k[x]$. O polinômio em duas variáveis $f(x) - f(y)$ é múltiplo de $x - y$, pois $x - y | (x^n - y^n)$ para todo $n \in \mathbb{N}$. Seja $g(x, y)$ o polinômio tal que $f(x) - f(y) = g(x, y) \cdot (x - y)$. Chamamos de **derivada de** f o polinômio $f'(x) \doteq g(x, x)$.

O procedimento usado nesta definição, se por um lado pode ser visto apenas como uma interpretação formal de como calcular a derivada de um polinômio, também é natural quando investigamos a existência de raízes múltiplas:

Proposição 3.1.2. *Sejam k e $f(x)$ como na definição acima (3.1.1). Um elemento $a \in k$ é tal que $(x - a)^2 | f(x)$ se, e somente se, $f(a) = f'(a) = 0$.*

Demonstração. A condição $f(a) = 0$ nada mais é que $(x - a) | f(x)$. Pondo $p(x) \doteq \frac{f(x)}{x-a}$, queremos mostrar que $p(a) = 0$ se, e somente se, $f'(a) = 0$. Ora, $p(x) = \frac{f(x)}{x-a} = \frac{f(x)-f(a)}{x-a} = g(x, a)$, onde g está definido como em 3.1.1. Logo $p(a) = g(a, a) = f'(a)$. ■

Também valem as clássicas fórmulas para soma e produto:

Proposição 3.1.3. *Sejam $f_1, f_2 \in k[x]$ e $\alpha \in k$. Então:*

$$(i)\ (\alpha f_1)' = \alpha f_1'; \qquad (ii)\ (f_1 + f_2)' = f_1' + f_2'; \qquad (iii)\ (f_1 f_2)' = f_1' f_2 + f_1 f_2'.$$

Demonstração. Novamente, vamos usar a notação da definição 3.1.1: sejam $g_1, g_2 \in k[x, y]$ tais que

$$f_1(x) = f_1(y) + g_1(x, y) \cdot (x - y) \quad \text{e} \quad f_2(x) = f_2(y) + g_2(x, y) \cdot (x - y).$$

Então:

$$\frac{\alpha f_1(x)-\alpha f_1(y)}{x-y} = \alpha g_1(x,y) \quad \text{e} \quad \frac{(f_1+f_2)(x)-(f_1+f_2)(y)}{x-y} = (g_1+g_2)(x,y),$$

bem como temos

$$\frac{(f_1\cdot f_2)(x)-(f_1\cdot f_2)(y)}{x-y} = g_1(x,y)\cdot f_2(y) + f_1(y)\cdot g_2(x,y) + (g_1\cdot g_2)(x,y)\cdot(x-y).$$

Trocando y por x, temos o resultado. ■

Corolário 3.1.4. *Se $f(x) = \sum_{i=0}^{n} a_i x^i \in k[x]$, então $f'(x) = \sum_{i=1}^{n} i \cdot a_i^{i-1}$.*

Demonstração. Pelos itens i) e ii) da proposição acima, basta vericar para $f(x) = x^n$. Se $n = 0$, então $\frac{1-1}{/}x - y = 0$, logo $1' = 0$. Se $n = 1$, $\frac{x-y}{x-y} = 1$, logo $x' = 1$. Se $n \geq 2$ e o resultado vale para $n-1$:

$$(x^n)' = 1\cdot x^{n-1} + x\cdot(x^{n-1})' = x^n + (n-1)\cdot x\cdot x^{n-2} = n\cdot x^{n-1}.$$ ■

Corolário 3.1.5 (Regra da Cadeia)**.** *Sejam $f, h \in k[x]$. Então $(h\circ f)'(x) = h'(f(x))\cdot f'(x)$.*

Demonstração. Exatamente a mesma prova do corolário anterior, mas usando f^n ao invés de x^n e ignorando o caso $n = 1$. ■

Notação 3.1.6. Denotamos por $f^{(n)}$ a *n-ésima derivada* de um polinômio f, ou seja, ao resultado de n derivações consecutivas do polinômio f. Recursivamente:

- $f^{(0)} = f$;
- $f^{(n)} = (f^{(n-1)})'$.

Proposição 3.1.7 (Polinômio de Taylor em torno da origem)**.** *Seja $f(x) \in k[x]$. Então*

$$f(x) = \sum_{n=0}^{+\infty} \frac{f^{(n)}(0)}{n!} x^n,$$

onde tal soma infinita é interpretada como a soma dos termos não-nulos, os quais ocorrem apenas em número finito.

Demonstração. Seja $f(x) = \sum_{i=0}^{n} a_i x^i$. Para cada $i \geq k$, afirmamos que o coeficiente do termo de grau $i-k$ de $f^{(k)}$ é $a_i \cdot \prod_{l=i-k+1}^{i} l$. Verifiquemos por indução em k; se $k = 0$, o coeficiente do termo de grau i de $f^{(0)}$ é a_i. Suponhamos o resultado verdadeiro para um certo k_0. Por 3.1.4, o termo de grau $i-(k_0+1)$ de $f^{(k_0+1)}$ é a derivada do termo de grau $i-k_0$ de $f^{(k_0)}$, ou seja, é $(i-k_0)\cdot a_i \cdot \prod_{l=i-k_0+1}^{i} l$. Em particular, no caso $k = i$, temos $a_i \cdot i! = f^{(i)}(0)$ e, no caso $k > n$, $f^{(k)} = 0$. ■

Corolário 3.1.8 (Polinômio de Taylor em torno de um ponto qualquer)**.** *Seja $f(x) \in k[x]$. Então*

$$f(x+a) = \sum_{n=0}^{+\infty} \frac{f^{(n)}(a)}{n!} x^n,$$

ou, equivalentemente,

$$f(x) = \sum_{n=0}^{+\infty} \frac{f^{(n)}(a)}{n!} (x-a)^n.$$

Demonstração. Segue diretamente de 3.1.5 e 3.1.7. ■

3.2 Derivadas em um corpo real fechado

No decorrer dessa seção, $\mathbf{R}$ é um corpo real fechado fixado (ainda que arbitrário) e $\leq$ sua única ordem.

Proposição 3.2.1 (Teorema de Rolle)**.** *Sejam $f \in \boldsymbol{R}[x]$ e $a, b \in \boldsymbol{R}$, com $a < b$. Se $f(a) = f(b)$ então existe $c \in (a, b)$ tal que $f'(c) = 0$.*

Demonstração. Seja $g(x) \doteq f(x) - f(b)$. Se g for nula, o resultado é trivial. Caso contrário, podemos assumir, sem perda de generalidade, que $a, b \in \mathbf{R}$ são tais que g nunca se anula em $[a, b]$.

Como $g(a) = g(b) = 0$, temos que $(x-a)(x-b)$ divide $g(x)$. Seja $h(x) \in \mathbf{R}[x]$ tal que $g(x) = (x-a)^n(x-b)^m h(x)$ e com $h(x)$ não se anulando em $[a, b]$. Mas então

$$f'(x) = g'(x) = n(x-a)^{n-1}(x-b)^m h(x) + m(x-a)^n(x-b)^{m-1}h(x) + (x-a)^n(x-b)^m h'(x)$$

$$= (x-a)^{n-1}(x-b)^{m-1}(\tilde{h}(x))$$

onde $\tilde{h}(x) = n(x-b)h(x) + m(x-a)h(x) + (x-a)(x-b)h'(x)$. Mas então $\tilde{h}(a) = n(a-b)h(a)$ e $\tilde{h}(b) = -m(a-b)h(b)$. Logo $\tilde{h}(x)$ muda de sinal em $[a, b]$, tendo uma raiz $c \in (a, b)$, a qual, com mais razão, é também raiz de $f'(x)$. ■

Corolário 3.2.2 (Teorema do Valor Médio)**.** *Sejam $f(x) \in \boldsymbol{R}[x]$ e $a, b \in \boldsymbol{R}$. Então existe $c \in (a, b)$ tal que $f'(c) = \frac{f(b)-f(a)}{b-a}$.*

Demonstração. Aqui podemos usar a demonstração usual dos cursos de cálculo. Definimos $g(x) \doteq f(x) - \frac{x-a}{b-a} \cdot (f(b) - f(a))$. Assim $g(a) = f(a) = g(b)$. Pela proposição anterior (3.2.1), existe $c \in (a, b)$ com $0 = g'(c) = f'(c) - \frac{f(b)-f(a)}{b-a}$, concluindo a demonstração. ■

Corolário 3.2.3. *Sejam R um corpo real fechado e $p(x) \in R[x]$. Se $c \in R$ é tal que $p'(c) > 0$, então existe $\delta > 0$ tal que $p(x)$ é estritamente crescente no intervalo $(c - \delta,\, c + \delta)$. Enunciado análogo vale no caso em que $p'(c) < 0$.*

Demonstração. Pela conservação de sinal (2.2.3), existe $\delta > 0$ tal que $p'(x)$ é estritamente positivo no intervalo $I = (c-\delta,\, c+\delta)$. Agora, o mesmo argumento dos cursos elementares de Cálculo, utilizando o Teorema do Valor Médio (3.2.2), mostra que p é estritamente crescente em I. Analogamente, trata-se o caso em que $p'(c) < 0$. ■

3.3 Contando Raízes

Embora encontrar as raízes de um polinômio possa ser bastante complicado, calcular *quantas* são as raizes é bem mais simples.

Vamos seguir nesta seção um caminho histórico que se seguiu para atacar esse problema. Começando com um resultado apresentado por Descartes no "Discurso do Método", que depois foi refinado por Budan e Fourier e chegando a uma resposta definitiva com Sturm. Todos esses resultados se baseiam em contar *trocas de sinais*.

Observação. No decorrer desta seção, $\mathbf{R}$ denotará um *corpo real fechado*.

Não existe definição universal sobre como se deve contar o número de trocas de sinal em uma sequência. Claro que se a e b são dois elementos adjacentes em uma sequencia com $ab < 0$, devemos contar uma troca, e se $ab > 0$ não. O problema é quando $ab = 0$. Assim, primeiramente vamos definir o *número de trocas de sinal* em uma sequência finita de elementos não-nulos de k, onde k é um corpo ordenado.

Definição 3.3.1. Seja k um corpo ordenado, $n \in \mathbb{N}$ e $\boldsymbol{x} \doteq (x_1, \cdots, x_n) \in (k^*)^n$. Definimos o *número de trocas de sinal de* $\boldsymbol{x}$, denotado por $V(\boldsymbol{x})$, por indução em n:

- Se $n \in \{0, 1\}$, $V(\boldsymbol{x}) \doteq 0$;
- Se $n > 1$, $V(\boldsymbol{x}) \doteq \begin{cases} V(x_1, \cdots, x_{n-1}) + 1 & \text{se } x_{n-1}x_n < 0 \\ V(x_1, \cdots, x_{n-1}) & \text{se } x_{n-1}x_n > 0 \end{cases}$

Para lidar com o zero, seguiremos o caminho mais padrão, simplesmente omitindo os zeros da sequência dada e contando as trocas de sinais da nova sequência.

Definição 3.3.2. Seja $\boldsymbol{x} \doteq (x_1, \cdots, x_n) \in k^n$. Chamemos de m o número de coordenadas não nulas de $\boldsymbol{x}$ e definimos $\boldsymbol{x'} \doteq (x_{i_1}, \cdots, x_{i_m}) \in k^m$ onde $x_{i_1}, \cdots, x_{i_m}$ são as coordenadas não nulas de $\boldsymbol{x}$ e $i_r < i_s$ se $r < s$.

Definimos $V(\boldsymbol{x}) \doteq V(\boldsymbol{x'})$, este último valor como definido em 3.3.1. Claramente, não há conflito nas definições.

Nosso primeiro resultado é atribuído a Descartes por estar enunciado, sem prova, no famoso apêndice *La Géométrie* do livro *Discours de la Méthode*.Entretanto há evidências de que outros matemáticos já o conheciam antes e demonstrações só vieram depois. A prova que apresentamos aqui foi essencialmente retirada de [2].

Teorema 3.3.3. [Regra dos sinais de Descartes] *Seja* $\boldsymbol{R}$ *um corpo real fechado e seja* $f = a_n x^n + \cdots + a_1 x + a_0 \in \boldsymbol{R}[x]$, $a_n \neq 0$. *Então o numero de raízes estritamente positivas* distintas *de* f *é menor ou igual a* $V(a_n, \cdots, a_0)$.

Demonstração. Para $n = 0$, o resultado é trivial. Suponhamos $n > 0$ e que o resultado vale para os polinômios com grau $n - 1$. Se $a_0 = 0$, o resultado vale pois todas as raízes estritamente positivas de f são raízes de $\frac{f}{x} = a_n x^{n-1} + \cdots + a_2 x + a_1$. Assim, nos importemos com o caso $a_0 \neq 0$.

Temos que resultado vale para $f' = na_n x^{n-1} + \cdots + 2a_2 x + a_1$. Sejam k o número de raízes distintas de f' e $c_1 < \cdots < c_k$ tais raízes. Definamos $c_{k+1} \doteq +\infty$.

O Teorema de Rolle (3.2.1) garante que em $[c_i, c_{i+1}[$, $1 \leq i \leq k$, há no máximo uma raiz de f e, portanto, k limita superiormente o número de raízes de f que são maiores ou iguais a c_1. Mas

$$k \leq V(na_n, (n-1)a_{n-1}, \cdots, a_1) = V(a_n, a_{n-1}, \cdots, a_1).$$

Resta saber quantas raízes f pode ter em $]0, c_1[$. Seja $q \in \{1, \cdots, n\}$ tal que $a_q \neq 0$ e $a_i = 0$ para todo i tal que $0 < i < q$. Assim, $f' = na_n x^{n-1} + \cdots + qa_q x^{q-1}$. O sinal de f' não muda em $]0, c_1[$ e tal sinal é mesmo sinal de $\frac{f'}{x^{q-1}}$ neste mesmo intervalo, o qual por sua vez é o sinal de a_q, pois este é o mesmo sinal de $\frac{f'}{x^{q-1}}(0) = qa_q$. Assim só é possível para f ter no máximo uma raiz neste intervalo e ainda assim em apenas em 2 casos: $a_0 > 0$ e f' decrescente em $]0, c[$ ou $a_0 < 0$ e f' crescente em $]0, c[$; em ambos temos $a_0 a_q < 0$, completando a prova. ■

Fourier refinou o resultado acima usando Polinômio de Taylor (3.1.8):

Corolário 3.3.4. *Sejam $f \in \boldsymbol{R}[x]$, com grau n, e $c \in \boldsymbol{R}$. Então, $V(f(c), f'(c), \cdots, f^{(n)}(c))$ limita o número de raízes de f maiores que c.*

Demonstração. Basta notar que o número de raízes maiores que c é o número de raízes positivas de $f(x+c)$, que por 3.1.8 é igual a $\frac{f^{(n)}(c)}{n!}x^n + \cdots + f'(c)x + f(c)$. ■

3.3.1 O Teorema de Sturm

Observação. Como anteriormente, **R** indica um corpo real fechado.

Estudando textos de Fourier, Sturm decidiu examinar sequências de polinômios, sem que sejam necessariamente $f, f', \ldots, f^{(n)}$, mas que fossem capazes de ser um instrumento para determinar o número de raízes de um polinômio em um intervalo $[a, b]$ de **R** . Deste projeto surgiu a seguinte

Definição 3.3.5. A uma sequência de polinômios $p_0, p_1, \ldots, p_m \in \mathbf{R}[x]$ é dado o nome de *sequência de Sturm* se relativa ao intervalo $[a, b]$ se ela satisfaz as seguintes propriedades:

(i) $p_0(a) \cdot p_0(b) \neq 0$;

(ii) O polinômio $p_m(x)$ não se anula em $[a, b]$;

(iii) Se, para $1 \leq i \leq m-1$ e $c \in [a, b]$, $p_i(c) = 0$, então $p_{i-1}(c)$ e $p_{i+1}(c)$ são não nulos e tem sinais opostos;

(iv) Se $p_0(c) = 0$ então $p_1(c)$ tem o mesmo sinal que $p_0'(c)$.

Teorema 3.3.6 (Sturm, versão abstrata)**.** *Seja $p_0, p_1, \ldots, p_m \in \boldsymbol{R}[x]$ uma sequência de Sturm relativa ao intervalo $[a, b]$. Seja $N(c)$ o número de trocas de sinais da sequência $p_0(c), p_1(c), \ldots, p_m(c) \in \boldsymbol{R}$. Então o número de raízes de p_0 (sem multiplicidades) em $[a, b]$ é, exatamente, $N(a) - N(b)$.*

Demonstração. Vamos investigar para quais valores de c o número $N(c)$ se modifica. Para que o sinal de algum dos polinômios da sequência se modifique, é necessário que algum deles se anule. Como existe apenas um número finito de pontos em que isso pode acontecer, podemos restringir nossa análise ao caso de um intervalo $[a, b]$ onde $c \in [a, b]$ é o único ponto onde isso acontece.

Primeiramente, vamos supor que $p_i(c) = 0$, com $1 \leq i \leq m-1$ e olhar para a sequência de sinais de p_{i-1}, p_i, p_{i+1}. Como tanto $p_{i-1}(c)$ e $p_{i+1}(c)$ não se anulam em c, pelo que estamos supondo no caso em que estamos tratando, p_{i-1} e p_{i+1} são não nulos e não trocam de sinal em $[a, b]$, além de terem sinais opostos. Assim a sequência de sinais de p_{i-1}, p_i, p_{i+1} ou se mantém a mesma em a e em b, ou ocorre uma das seguintes mudanças: vai de $+, +, -$ para $+, -, -$, ou vai de $+, -, -$ para $+, +, -$, ou ainda os mesmos casos trocando os sinais. De qualquer forma, o número de troca de sinais permanece o mesmo!

Notemos que, pela definição de sequência de Sturm, não pode ser o caso em que $p_m(c) = 0$. Nos resta então verificar o que acontece quando $p_0(c) = 0$. Neste caso o item (iv) da definição nos resolve o problema: se $p'(c) > 0$, então temos que p_0 é crescente em $[a, b]$, logo $p_0(a) < 0$ e $p_0(b) > 0$, de onde concluímos que a sequência de sinais perde uma única troca. O raciocínio é análogo cuida do caso $p'(c) < 0$, completando a prova. ■

Exemplo 3.3.7. Se $f(x) = \frac{x^n}{n!} + \frac{x^{n-1}}{(n-1)!} + \cdots x + 1$, a sequência $f(x), f'(x), (-1)^{n+1}$ é de Sturm.

O Teorema 3.3.6 não nos diz muita coisa se não conseguirmos ter uma sequencia de Sturm em mãos. O ponto fundamental é que Sturm encontrou uma maneira de sempre calcular uma tal sequência.

Se queremos encontrar as raízes de um polinômio $p \in \mathbf{R}[x]$, claro que queremos colocar $p_0 = p$. Pela condição (iv), é natural colocarmos $p_1 = p'$. Mas o resto da sequência será obtido de uma maneira diferente, para garantir as demais condições.

Pelas condições exigidas, vamos ter que nos restringir a um polinômio $p \in \mathbf{R}[x]$ *sem* raízes múltiplas, pois $p'(c) \neq 0$ quando $p(c) = 0$. Essa não é uma restrição muito forte, pois sempre podemos "tornar" p sem raízes múltiplas usando o Algoritmo de Euclides.

Além disso, este algoritmo nos dá a chave de como completar o problema! A sequência de polinômios que obtemos por ele satisfaz (i), (ii) e (iv). Para conseguirmos (iii), basta fazer as seguintes modificações:

1. Trocar o sinal do resto!

2. Para garantir que os polinômios adjacentes não se anulem, basta dividir todos pelo último.

Definição 3.3.8. Seja f um polinômio em $\mathbf{R}[x]$. A sequência de polinômios $\{p_i\}_{0 \leq i \leq m}$ definida abaixo denomina-se a sequência de Sturm *canônica* de f:

(i) $p_0 = f$ e $p_1 = f'$;

(ii) Se $p_{n+1} \neq 0$, então p_{n+2} é o único polinômio tal que $p_n = q \cdot p_{n+1} - p_{n+2}$, com p_{n+2} nulo ou com grau menor que o grau de p_{n+1} ;

(iii) O último polinômio, p_m, é constante, mas o anterior, p_{m-1} não é.

Teorema 3.3.9 (Sturm, versão concreta)**.** *Seja $p_0 = f, p_1 = f', \ldots, p_m \in \boldsymbol{R}[x]$ a sequência canônica de Sturm de f . Seja $N(c)$ o número de trocas de sinais da sequência $p_0(c), p_1(c), \ldots, p_m(c) \in \boldsymbol{R}$. Então o número de raízes de p_0 (sem multiplicidades) em $[a, b]$ é, exatamente, $N(a) - N(b)$.*

Demonstração. Se $mdc(f, f') = 1$, então o a sequência de Sturm canônica de f é, de fato, uma sequência de Sturm: por hipótese f não se anula nos extremos do intervalo, p_m é uma constante não nula exatamente porque $mdc(f, f') = 1$, p_1, obviamente, tem sempre o sinal de f' e se $p_n(c) = 0$, $p_{n-1}(c) = -p_{n+1}(c)$, sendo que o $mdc(p_{n-1}, p_n)$ também é 1, logo $p_{n-1}(c) \neq 0$.

Se $mdc(f, f') = d$ com d não constante, então a sequencia de Sturm canônica não é uma sequência de Sturm, como se deveria esperar. Mas a sequência auxiliar $g_i = \frac{p_i}{d}$, com $i \in \{0, \ldots, m-1\}$ mantém as mesmas trocas de sinal em a e b (porque $d(a) \neq 0 \neq d(b)$), as raízes de $\frac{p}{d}$ são as mesmas de p e a nova sequência é de Sturm! O polinômio $\frac{f}{d}$ não se anula nos extremos do intervalo, o polinômio $\frac{p_{m-1}}{d}$ é uma constante não nula, e, se $f(c) = 0$, então temos

$$\left(\tfrac{f}{d}\right)'(c) = \tfrac{f'(c)d(c)+f(c)d'(c)}{d(c)^2} = \tfrac{f'(c)}{d(c)} = g_1(c),$$

completando a prova. ■

Como se pode notar, o Teorema de Sturm (3.3.9) pode ser usado para construir uma combinação booleana para decidir quantas raízes uma equação polinomial possui a partir dos seus coeficientes.

Basicamente estamos eliminando a variável x e o quantificador $\exists$. Como pode ser notado também, em geral isso dá muito trabalho. Utilizando o Teorema de Sturm, vamos enunciar e esboçar a prova de uma primeira versão, polinomial e algorítmica, do Teorema de Tarski-Seidenberg.

Teorema 3.3.10. *Existe um algoritmo tal que dado um polinômio $f \in \boldsymbol{R}[t_1, \ldots, t_n, X]$ produz uma lista finita de polinômios $R_1, \ldots, R_k \in \boldsymbol{R}[t_1, \ldots, t_n]$ e uma função $c : \{-1, 0, 1\}^k \longrightarrow \mathbb{N} \cup \{+\infty\}$ tais que, para todo $\epsilon = (\epsilon_1, \ldots, \epsilon_k) \in \{-1, 0, 1\}^k$ e toda n-upla $(t_1, \ldots, t_n) \in \boldsymbol{R}^n$, se*

$$sinal(R_1(t_1, \ldots, t_n)) = \epsilon_1, \ldots, sinal(R_k(t_1, \ldots, t_n)) = \epsilon_k,$$

então a equação em X, $f = (t_1, \ldots, t_n, X) = 0$, tem exatamente $c(\epsilon)$ raízes.

Em outras palavras, a quantidade de raízes (em $\mathbf{R}$) da equação $P(t_1, \ldots, t_n, X) = 0$ depende dos sinais de uma certa lista de polinômios em $t_1, \ldots, t_n$.

Esboço da demonstração do Teorema 3.3.10. Escrevemos o polinômio f com os coeficientes no corpo $\mathbf{R}(t_1, \ldots, t_n)$ e X sendo a variável. Primeiramente quebramos em três casos: os caso de o coeficiente do termo de maior grau, $R_1(t_1, \ldots, t_n)$, é igual a zero, maior que zero e o caso menor que zero. No caso de supormos igual a zero, aplicamos o mesmo algoritmo para o polinômio f menos o termo de maior grau.

No caso de considerarmos R_1 sendo diferente de zero, aplicamos o processo do teorema 3.3.9, sendo que após cada divisão quebramos a sequência canônica de Sturm em três: aquela em que o coeficiente de maior grau é nulo, aquela em que é positivo e aquela em que é negativo. A cada um desses coeficientes, multiplicado pelo quadrado do denominador para torná-lo um polinômio e manter o sinal, damos o nome de R_i para algum i ainda não usado. No caso de supormos R_i nulo, repetimos sempre o teste para o "novo" termo dominante.

Após isso, teremos diversas sequências de Sturm, cada uma associada a uma atribuição de sinais para cada R_i. Ora, mas sabendo os sinais dos coeficientes do termos dominantes de cada polinômio em uma sequência de Sturm conseguimos calcular os sinais de cada um deles em $-\infty$ e em $+\infty$, logo conseguimos calcular quantas raízes reais o polinômio f terá em cada caso. ■

O resultado acima pode ser facilmente generalizado para **sistemas** de equações polinomiais:

Corolário 3.3.11. *Dados $f_1, \ldots, f_k$ polinômios em $\boldsymbol{R}[t_1, \ldots, t_n, x]$, existe um algoritmo produz uma lista finita de polinômios $R_1, \ldots, R_k \in \boldsymbol{R}[t_1, \ldots, t_n]$ e uma função $c : \{-1, 0, 1\}^k \longrightarrow \mathbb{N} \cup \{+\infty\}$ tais que, para todo $\epsilon = (\epsilon_1, \ldots, \epsilon_k) \in \{-1, 0, 1\}^k$ e toda n-upla $(t_1, \ldots, t_n) \in \boldsymbol{R}^n$, se*

$$sinal(R_1(t_1, \ldots, t_n)) = \epsilon_1, \ldots, sinal(R_k(t_1, \ldots, t_n)) = \epsilon_k,$$

então o **sistema** *em X, $f_1(t_1, \ldots, t_n, X) = \cdots = f_k(t_1, \ldots, t_n, X) = 0$, tem exatamente $c(\epsilon)$ raízes.*

Demonstração. Basta aplicar o teorema 3.3.10 ao polinômio $f = f_1^2 + \cdots + f_k^2$. ■

3.4 Fecho Real de um Corpo Ordenado

Nesta seção discutiremos a existência e unicidade do fecho real de um corpo ordenado, utilizando idéias originárias de Serge Lang.

Definição 3.4.1. Um **fecho real** de um corpo ordenado, $\langle k, \leq \rangle$, é um corpo real fechado, R, tal que R é uma extensão algébrica de k e a ordem em k é a induzida pela (única) ordem em R. [1]

[1] Alguns autores definem fecho real exigindo apenas que seja uma extensão algébrica e real fechada do corpo base. Preferimos ir direto ao ponto...

O resultado fundamental acerca de fechos reais é o seguinte

Teorema 3.4.2. *Todo corpo ordenado possui um fecho real, que é único a menos de isomorphismo de corpos ordenados.*

Demonstração. Seja $\langle k, \leq \rangle$ um corpo ordenado e $\bar{k}$ o seu fecho algébrico. A prova será fruto de uma série de fatos.

Existência do fecho real. Começamos pelo seguinte Fato (sugerido pela construção do fecho pitagórico de um corpo formalmente real):

Fato 1. *Se K é o corpo gerado, dentro de $\bar{k}$, adicionando-se a k a raiz quadrada de todos os seus elementos positivos, então K é formalmente real.*

Prova. Suponha que não; então existem $c_1, \ldots, c_n \in K$ tais que

$$-1 = \sum_{j=1}^{n} c_j^2. \tag{I}$$

Seja ℓ o menor inteiro tal que existem c_k, $1 \leq k \leq n$, em uma extensão da forma $k(\sqrt{d_1}, \ldots, \sqrt{d_{\ell+1}})$, com $d_1, \ldots, d_{\ell+1} > 0$ em k e satisfazendo (I). Note que, pela minimalidade de ℓ, $\sqrt{d_{\ell+1}}$ não está em $k(\sqrt{d_1}, \ldots, \sqrt{d_\ell})$. Para cada $1 \leq j \leq n$, podemos escrever

$$c_j = u_j + v_j\sqrt{d_{\ell+1}}, \quad \text{com } u_j, v_j \in k(\sqrt{d_1}, \ldots, \sqrt{d_\ell}),$$

e portanto, de (I) vem

$$\begin{aligned} -1 &= \sum_{j=1}^{n} (u_j + v_j\sqrt{d_{\ell+1}})^2 = \sum_{j=1}^{n} u_j^2 + 2u_jv_j\sqrt{d_{\ell+1}} + v_j^2 d_{\ell+1} \\ &= \sum_{j=1}^{n} u_j^2 + v_j^2 d_{\ell+1} + \sqrt{d_{\ell+1}} \sum_{j=1}^{n} 2u_jv_j. \end{aligned} \tag{II}$$

Já que $\sqrt{d_{\ell+1}}$ não pode estar em $k(\sqrt{d_1}, \ldots, \sqrt{d_\ell})$, (II) acarreta $-1 = \sum_{j=1}^{n} u_j^2 + v_j^2 d_{\ell+1}$, contrariando a minimalidade de ℓ e encerrando a prova do Fato 1. □

Agora, exatamente como no início da prova do Teorema 2.4.10, o Lema de Zorn fornece uma extensão algébrica e real fechada, R, da extensão algébrica K construída no Fato 1. Então, R é um fecho real de k. De fato, basta mostrar que a ordem de k é a induzida pela de R. Mas note que para todo $d > 0$ em k, d é um quadrado distinto de zero em R, e portanto $d > 0$ na ordem de R; recipricamente, se $d \in k$ é tal que $d >_R 0$ e $d <_k 0$, então também teríamos $-d >_R$, uma contradição. Logo, a ordem de k é a induzida pela de R, completando a prova da existência do fecho real de $\langle k, \leq \rangle$.

Unicidade a menos de isomorfismo. Sejam R e R' fechos reais de $\langle k, \leq \rangle$. Podemos supor que $k \neq R$, R', senão, nada a provar (um corpo real fechado não possui extensão algébrica própria e formamente real).

A estratégia é usar o Lema de Zorn. Para isto, precisamos primeiro mostrar:

(*) Se K é um subcorpo próprio de R contendo k, $\tau : K \longrightarrow R'$ é uma imersão de corpos ordenados sobre k e $\alpha \in R \setminus K$, então existe um par $\langle F, \sigma \rangle$ onde F é uma extensão finita de K contendo α e $\sigma : F \longrightarrow K'$ é uma imersão de corpos ordenados que estende τ.

Iniciamos com o seguinte

Fato 2. *Seja $\langle L, \leq \rangle$ um corpo ordenado e M uma extensão algébrica e ordenada de L, que induz em L a ordem original. Então, para todo $u >_M 0$ em M, existe $v >_L 0$ em L, tal que $v <_M u$.* [2]

[2] Ou seja, M não possui infinitesimais em relação a L.

Prova do Fato 2. Suponha resultado falso, e seja $u >_M 0$ em M, tal que $u <_M z$, para todo $z \in L$. É fácil ver que para todo $\alpha \neq 0$ em L e todo $n \geq 1$ em $\mathbb{N}$, $|\alpha|u^k <_M z$, qualquer que seja $z >_k 0$ em L. Mas então, a desigualdade triangular (2.2.3) garante que se $f(x) \in L[x]$ é um polinômio de termo constante 0, temos $|f(u)| <_M z$, para todo $z \in L$.

Como M é algébrico sobre L, seja $p(x) = xq(x) + z_0 \in L[x]$ o polinômio minimal (e irredutível) de u sobre L. Então de $p(u) = 0$, obtemos $uq(u) = -z_0$, ou seja, $|uq(u)| = |z_0|$, uma contradição. □

Seja $\langle K, \tau\rangle$ um par como acima e $\alpha \in R \setminus K$. Indicaremos por K' a imagem τ em R'.

Seja $p(x) \in K[x]$ o polinômio minimal de α sobre K; escrevemos $p^\tau(x)$ para o polinômio cujos coeficientes são a imagem pela imersão τ. Como estamos em característica zero, sabemos que $p(\alpha) = 0$, mas $p'(\alpha) \neq 0$. Pelo corolário 3.2.3, existe $\delta_1 <_R 0$ tal que o polinômia $p(\alpha + h)$, $h \in K$, é estritamente crescente ou estritamente descrescente no intervalo $(\alpha-\delta_1, \alpha+\delta_1)$; em particular, para $h \in K$, $p(\alpha + h)$ troca de sinal neste intervalo. Pelo Fato 2, existe $\delta >_K 0$ em K tal que $\delta <_R \delta'$ e portanto $p(\alpha + h)$ também troca de sinal no intervalo $I = (\alpha - \delta, \alpha + \delta)$. Uma nova aplicação do Fato 2 fornece $u \in K$, tal que $|\alpha - u| < \frac{\delta}{4}$ e assim o intervalo $J = (u - \frac{\delta}{4}, u + \frac{\delta}{4}) \subseteq I$. Logo, o polinômio $p(u + h)$, $h \in K$, troca de sinal em J. Concluímos que o polinômio $\tau(p(u + h)$ troca de sinal no intervalo $\tau(J)$ e o teorema 2.3.6 aplicado a R' garante que existe $\beta \in R'$ tal que $p^\tau(\beta) = 0$. Portanto, existe um isomorfismo de $K(\alpha)$ em $K'(\beta)$, estendendo τ.

Fato 3. *Com a notação acima, sejam*

$$\begin{cases} r_1 <_R r_2 <_R \cdots <_R r_n & \text{as raízes distintas de } p \text{ em } R; \\ s_1 <_{R'} s_2 <'_R \cdots <_{R'} s_m & \text{as raízes distintas de } p^\tau \text{ em } R'. \end{cases}$$

Então, $n = m$ e existe uma imersão σ de corpos ordenados de $K(r_1, \ldots, r_n)$ em R' que é extensão de τ e satisfaz $\sigma(r_j) = s_j$, $1 \leq j \leq n$.

Prova do Fato 3. Para $1 \leq j \leq n$, seja $u_j \in R$ tal que $u_j^2 = r_{j+1} - r_j$ $(1 \leq j \leq (n-1))$. Seja $K_1 = K(r_1, \ldots, r_n, u_1, \ldots, u_{n-1})$; pelo estabelecido acima, existe uma imersão (de corpos),

$$\tau' : K_1 \longrightarrow R',$$

estendendo τ. Note que $\tau'(r_{j+1}) - \tau'(r_j) = \tau'(u_j^2)$ é um quadrado em R' e portanto, temos

$$\tau'(r_1) <_{R'} \ldots <_{R'} \tau'(r_n) \text{ em } R'.$$

Concluímos imediatamente que $n \leq m$ e, por simetria, que $n = m$.

Seja $\sigma : F = K[r_1, \ldots, r_n] \longrightarrow R'$ a extensão de τ determinada pelas condições $\sigma(r_j) = s_j$, $1 \leq j \leq n$. Resta provar que σ preserva ordem. Seja $v \in F$, com $v >_R 0$. Então, existe $u \in R$ tal que $u^2 = v$. Como acima, existe uma imersão τ^*, de $F[u_1, \ldots, u_{n-1}, u]$ em R', estendendo σ e tal que $\sigma(v) = \tau^*(u^2)$, i.e., $\sigma(v) >_{R'} 0$, completando a prova do Fato 3. □

A afirmação em (*) acima é consequência imediata do Fato 3. Agora, uma aplicação do Lema de Zorn fornece uma imersão de corpos ordenados sobre k, $\eta : R \longrightarrow R'$. Que η é sobrejetora segue da simetria do argumento, encerrando a prova do teorema 3.4.2. ■

Observação. Na realidade, se R e R' são fechos reais de k, o isomorfismo do Teorema 3.4.2 é **único**. Deixamos este fato, que não será usado mais tarde, aos cuidados do leitor.

Capítulo 4

Teoria dos Modelos e Aplicações

Neste capítulo vamos admitir que o leitor tenha familiaridade com os conceitos básicos de linguagens de primeira ordem com igualdade, suas interpretações, a noção de satisfação, bem como a de modelo de uma teoria. Uma referência standard para este tema é [3].

Se $\mathcal{L}$ é uma linguagem de primeira ordem com igualdade, lembramos que uma formula φ de $\mathcal{L}$ é
• uma **sentença** se não tiver variáveis livres;
• **sem quantificadores** (ou livre de quantificadores) se não houver ocorrência dos quantificadores $\exists$ e $\forall$ em φ.
Uma **teoria** em $\mathcal{L}$ é um conjunto de sentenças de $\mathcal{L}$. Se M é uma $\mathcal{L}$-estrutura e T é uma teoria em $\mathcal{L}$, como é usual, escrevemos $M \models T$ se todo elemento de T está satisfeito em M.

Fundamentais para nós serão duas particulares linguagens de primeira ordem com igualdade: a dos *anéis* e a dos *anéis ordenados*, indicadas, respectivamente, por

$$\mathcal{L}_a = \langle +, \cdot, 0, 1\rangle \quad \text{e} \quad \mathcal{L}_{ao} = \langle +, \cdot, 0, 1, \leq\rangle.$$

4.1 Conjuntos Semialgébricos Revisitados

As linguagens de primeira ordem nos permitem reescrever a definição de conjunto semialgébrico de uma nova maneira:

Proposição 4.1.1. *Seja R um corpo ordenado. Um subconjunto S de R^n é semialgébrico se, e somente se, existe uma fórmula sem quantificadores da linguagem dos anéis ordenados, P, nas variáveis livres $x_1, \ldots, x_n$ tal que $S = \{(a_1, \ldots, a_n) \in R^n \mid R \models P(a_1, \ldots, a_n)\}$.*

Demonstração. A demonstração desse fato é simples. Os pontos menos claros são como traduzir a negação e a implicação em uniões e intersecções. A implicação é possível se soubermos a negação. Para a negação nós vamos usar as relações de DeMorgan e o que sabemos sobre uma relação de ordem total: $a \neq b \iff a < b$ ou $b < a$ e $\neg(a < b) \iff b < a$ ou $b = a$. ■

A projeção de um conjunto semialgébrico, entretanto, não é obviamente descritível desta forma. Se $S = \{(a_1, \ldots, a_n) \in R^n \mid R \models P(a_1, \ldots, a_n)\}$ então a projeção nas $m < n$ primeiras coordenadas é o conjunto

$$\{(a_1, \ldots, a_m) \in R^m \mid \text{ existem } a_{m+1}, \ldots, a_n \text{ para os quais } R \models P(a_1, \ldots, a_m, a_{m+1}, \ldots a_n)\}.$$

A versão geométrica do Teorema de Tarski-Seidenberg (2.1.3) nos diz que, no caso dos corpos reais fechados, há um modo de nos livrar desse "existem"...

4.2 Eliminação de Quantificadores

Nessa seção veremos técnicas de como provar que uma teoria admite *Eliminação de Quantificadores*, para aplicarmos posteriormente no exemplo das teorias dos *Corpos Algebricamente fechados* e dos *Corpos Reais Fechados*.

O próximo resultado afirma que a validade de uma afirmação sem quantificadores é local, isto é, depende apenas dos elementos envolvidos na afirmação, sem ter que fazer uma busca por outros elementos (como exigem o $\forall$ e o $\exists$).

Proposição 4.2.1. *Seja $\mathcal{L}$ uma linguagem de primeira ordem com igualdade e $\mathcal{M}$, $\mathcal{N}$ $\mathcal{L}$-estruturas. Se $\mathcal{M}$ é uma subestrura de $\mathcal{N}$, $a_1, \ldots, a_n \in M$ e $\phi(x_1, \ldots, x_n)$ é uma fórmula sem quantificadores, então*

$$\mathcal{M} \models \phi(a_1, \ldots, a_n) \quad \Leftrightarrow \mathcal{N} \models \phi(a_1, \ldots, a_n).$$

Demonstração. Primeiramente temos que, na verdade, se $t(x_1, \ldots, x_n)$ é um termo, então sua interpretação é a mesma nas duas estruturas, ou melhor dizendo, se $b_1, \ldots, b_n \in M$, então $t^{\mathcal{M}}(b_1, \ldots, b_n) = t^{\mathcal{N}}(b_1, \ldots, b_n)$. Provemos isso por indução nos termos:

- Se t é uma constante c, então $c^{\mathcal{M}} = c^{\mathcal{N}}$;
- Se t é a variável x_i, $t^{\mathcal{M}}(b_1, \ldots, b_n) = b_i = t^{\mathcal{N}}(b_1, \ldots, b_n)$;
- Se já temos o resultado para os termos $t_i, \ldots, t_n$ e f é uma operação n-ária, então, como a intepretação da função e $\mathcal{M}$ é a restrição da interpretação em $\mathcal{N}$, ao "calcularmos" a função nos termos e substituirmos as variáveis por $b_1, \ldots, b_n$ teremos o mesmo resultado em ambos modelos.

Agora vamos para o resultado em si, também por indução, mas agora na complexidade das fórmulas.

• Primeiramente, vale para as atômicas. Seja R é uma relação n-ária:

$$\mathcal{M} \models \phi(a_1, \ldots, a_n) \Leftrightarrow (t_1(a_1, \ldots, a_n), \ldots, t_n(a_1, \ldots, a_n)) \in R^{\mathcal{M}}$$
$$\Leftrightarrow (t_1(a_1, \ldots, a_n), \ldots, t_n(a_1, \ldots, a_n)) \in R^{\mathcal{N}} \Leftrightarrow \mathcal{N} \models \phi(a_1, \ldots, a_n).$$

• Para a negação:

$$\mathcal{M} \models \neg\phi(a_1, \ldots, a_n) \Leftrightarrow \mathcal{M} \not\models \phi(a_1, \ldots, a_n) \Leftrightarrow \mathcal{N} \not\models \phi(a_1, \ldots, a_n) \Leftrightarrow \mathcal{N} \models \neg\phi(a_1, \ldots, a_n).$$

• Conjunção:

$$\mathcal{M} \models \phi \wedge \psi \Leftrightarrow \mathcal{M} \models \phi \text{ e } \mathcal{M} \models \psi \Leftrightarrow \mathcal{N} \models \phi \text{ e } \mathcal{N} \models \psi \Leftrightarrow \mathcal{N} \models \phi \wedge \psi.$$

• As regras de *DeMorgan* completam as outras (implicação e disjunção). ■

Definição 4.2.2. Dizemos que uma teoria T admite *Eliminação de Quantificadores* se para toda fórmula $\phi(x_1, \ldots, x_n)$ existe uma fórmula sem quantificadores $\psi(x_1, \ldots, x_n)$ tal que $T \models (\forall x_1, \ldots, \forall x_n)(\phi(x_1, \ldots, x_n) \leftrightarrow \psi(x_1, \ldots, x_n))$.

Nosso objetivo é caracterizar a Eliminação de Quantificadores por meio de modelos da teoria. Esse trabalho é feito pelo seguinte teorema:

Teorema 4.2.3. *Sejam $\mathcal{L}$ uma linguagem de primeira ordem com um símbolo constante c, T uma $\mathcal{L}$-teoria e $\phi(x_1, \ldots, x_n)$ uma $\mathcal{L}$-fórmula. São equivalentes:*

(i) Existe uma fórmula sem quantificadores $\psi(x_1, \ldots, x_n)$ tal que

$$T \models (\forall x_1, \ldots, \forall x_n)(\phi(x_1, \ldots, x_n) \leftrightarrow \psi(x_1, \ldots, x_n));$$

(ii) Se $\mathcal{M}$ e $\mathcal{N}$ são modelos de T e A é uma subestrutura comum de ambos, então $\mathcal{M} \models \phi(a_1, \ldots, a_n)$ se, e somente se, $\mathcal{N} \models \phi(a_1, \ldots, a_n)$, para toda sequência $a_1, \ldots, a_n \in A$.

A primeira implicação é exatamente a proposição 4.2.1. Para demonstrar $(ii) \to (i)$, precisamos de mais duas definições e de um lema:

Definição 4.2.4. Seja $\mathcal{M}$ uma $\mathcal{L}$-estrutura. Chamamos de $\mathcal{L}_M$ a linguagem obtida adicionando-se um novo símbolo de constante, m, para cada $m \in \mathcal{M}$.

Definição 4.2.5. O *Diagrama atômico de M*, representado por $Diag(M)$, é o conjunto

$$\{\phi(m_1, \ldots, m_n) \mid \phi \text{ é atômica ou negação de atômica e } M \models \phi(m_1, \ldots, m_n)\}$$

Lema 4.2.6. *Seja $\mathcal{N}$ uma $\mathcal{L}_M$ estrutura tal que $\mathcal{N} \models Diag(M)$. Então existe um $\mathcal{L}$-morfismo injetor de $\mathcal{M}$ em $\mathcal{N}$.*

Demonstração. Basta definir f como esperado: a todo elemento $m \in \mathcal{M}$ levamos a interpretação da constante $m^{\mathcal{N}}$ em $\mathcal{N}$. Certamente é injetiva, pois se $m_1 \neq m_2$ em $\mathcal{M}$, $(m_1 \neq m_2) \in Diag(\mathcal{M})$, logo $f(m_1) \neq f(m_2)$. Podemos proceder analogamente para vermos que as funções e as relações de $\mathcal{L}$ também são preservadas. ■

Agora vamos a demonstração da parte que falta de 4.2.3:

Demonstração. Vamos fazer essa demonstração em partes para melhor compreensão.
- Se $T \models (\forall x_1, \ldots, \forall x_n)\phi(x_1, \ldots, x_n)$, então $T \models (\forall x_1, \ldots, \forall x_n)(\phi(x_1, \ldots, x_n) \leftrightarrow c = c)$;
- Se $T \models (\forall x_1, \ldots, \forall x_n)\neg\phi(x_1, \ldots, x_n)$, então $T \models (\forall x_1, \ldots, \forall x_n)(\phi(x_1, \ldots, x_n) \leftrightarrow c \neq c)$.

Sendo assim, podemos supor que tanto $T \cup \{\phi(x_1, \ldots, x_n)\}$ quanto $T \cup \{\neg\phi(x_1, \ldots, x_n)\}$ são satisfazíveis no que se segue.

A estratégia consiste em pegar todas as consequências livres de quantificadores de ϕ e mostrar que existe um quantidade *finita* (e portanto uma única, tomando a conjunção) que implica ϕ.

Seja

$$\Gamma(x_1, \ldots, x_n) = \{\psi(x_1, \ldots, x_n) \mid \psi \text{ não tem quantificadores e } T \models (\forall x_1, \ldots, \forall x_n)(\phi(x_1, \ldots, x_n) \to \psi(x_1, \ldots, x_n)\},$$

e incorporemos $d_1, \ldots, d_m$ como novas constantes da linguagem. Vamos supor que vale que

$$T \cup \Gamma(d_1, \ldots, d_n) \models \phi(d_1, \ldots, d_n). \tag{4.1}$$

Isso sendo verdade, temos que $T \cup \Gamma(d_1, \ldots, d_n) \cup \{\neg\phi(d_1, \ldots, d_n)\}$ é insatisfazível. Usando o teorema da compacidade (4.3.11), temos que existem $\psi_1(x_1, \ldots, x_n), \ldots, \psi_k(x_1, \ldots, x_n) \in \Gamma(x_1, \ldots, x_n)$ tais que $T \cup \{\psi_1(d_1, \ldots, d_n), \ldots, \psi_k(d_1, \ldots, d_n)\} \cup \{\neg\phi(d_1, \ldots, d_n)\}$ é insatisfazível, de onde concluímos que

$$T \cup \{\psi_1(d_1, \ldots, d_n), \ldots, \psi_k(d_1, \ldots, d_n)\} \models \phi(d_1, \ldots, d_n),$$

o que é equivalente a

$$T \models (\bigwedge_{i=1}^{k} \varphi_i(d_1, \ldots, d_n) \to \phi(d_1, \ldots, d_n)).$$

Mas como T não nos restringe em nada as atribuições para os símbolos de constante d_i, obtemos

$$T \models (\forall x_1, \ldots, \forall x_n)(\bigwedge_{i=1}^{k} \varphi_i(x_1, \ldots, x_n) \to \phi(x_1, \ldots, x_n)),$$

de onde concluímos que ϕ é equivalente a uma fórmula sem quantificadores.

Isso reduz o problema a demonstrar que vale 4.1.

Suponha que não e seja $\mathcal{M} \models T \cup \Gamma(d_1, \ldots, d_n) \cup \{\neg\phi(d_1, \ldots, d_n)\}$. Chamemos de A a menor subestrutura de $\mathcal{M}$ que contém $d_1, \ldots, d_n$ (existe pois intersecção de subestruturas é subestrutura).

Para a existência de $\mathcal{N}$, precisamos que $\mathcal{A} \subseteq \mathcal{N}$. Para conseguirmos tal feito, é necessário e suficiente (usando o lema 4.2.6) que $\Sigma = T \cup Diag(\mathcal{A}) \cup \{\phi(d_1, \ldots, d_n)\}$ seja satisfazível.

Suponhamos que não; por compacidade, existiriam $\varphi_1(d_1, \ldots, d_n), \ldots, \varphi_r(d_1, \ldots, d_n) \in Diag(\mathcal{A})$ (logo sem quantificadores) tais que, por argumento análogo ao de antes,

$$T \models (\forall x_1, \ldots, \forall x_n)(\bigwedge_{i=1}^{r} \varphi_i(x_1, \ldots, x_n) \to \neg\phi(x_1, \ldots, x_n)),$$

e portanto

$$T \models (\forall x_1, \ldots, \forall x_n)(\phi(x_1, \ldots, x_n) \to \bigvee_{i=1}^{r} \neg\varphi_i(x_1, \ldots, x_n))$$

e assim $\bigvee_{i=1}^{r} \neg\varphi_i(x_1, \ldots, x_n) \in \Gamma(x_1, \ldots, x_n)$, e sendo $\mathcal{A}$ subestrutura de $\mathcal{M}$, $\mathcal{A} \models \bigvee_{i=1}^{r} \neg\varphi_i(d_1, \ldots, d_n)$, um absurdo, pois cada $\varphi_i(d_1, \ldots, d_n)$ está em $Diag(\mathcal{A})$. ■

O próximo resultado nos diz que podemos retirar apenas um quantificador de cada vez. A demonstração é simplesmente uma indução na complexidade das fórmulas e, portanto, será omitida.

Lema 4.2.7. *Seja T uma $\mathcal{L}$-teoria. Se para toda fórmula sem quantificadores $\theta(x_1, \ldots, x_n, w)$ existe uma fórmula também sem quantificadores $\phi(x_1, \ldots, x_n)$ tal que*

$$T \models (\forall x_1, \ldots, \forall x_n)((\exists w)\theta(x_1, \ldots, x_n, w) \leftrightarrow \phi(x_1, \ldots, x_n)),$$

então T admite eliminação de quantificadores.

Juntando este lema com o teorema 4.2.3, é imediato o seguinte resultado, que nos será útil no que segue:

Corolário 4.2.8. *Sejam $\mathcal{M}$ e $\mathcal{N}$ $\mathcal{L}$-estruturas, $\mathcal{M}, \mathcal{N} \models T$, A uma sub-estrutura de ambos, $a_1, \ldots, a_n \in A$, e $\phi(x_1, \ldots, x_n, w)$ uma fórmula sem quantificadores. Se vale*

$$\mathcal{M} \models (\exists w)\phi(x_1, \ldots, x_n, w) \iff \mathcal{N} \models (\exists w)\phi(a_1, \ldots, a_n, w),$$

então T admite Eliminação de Quantificadores.

4.2.1 Corpos Reais Fechados

Nesse ponto vale a pena lembrar que existe uma ambiguidade: na verdade temos duas teorias dos corpos reais fechados, uma na linguagem dos anéis e outra na linguagem dos anéis ordenados. Nós vamos ver que a primeira não admite eliminação de quantificadores enquanto a segunda, sim, admite.

Teoria dos Corpos Reais Fechados na linguagem dos anéis ordenados

Teorema 4.2.9. (Tarski-Seindenberg) *A teoria dos corpos reais fechados admite eliminação de quantificadores.*

Demonstração. Sejam K e L corpos reais fechados e $A \subset K \cap L$ uma subestrutura comum de ambos. Como intersecção de corpos ordenados é um corpo ordenado, temos um corpo ordenado contido em ambos K e L. Chamemos de F o fecho real de $K \cap L$ dentro de L (que nada mais é do que o conjunto dos elementos de K que são algébricos sobre $K \cap L$).

Seja $\phi(x_1, \ldots, x_n, y)$ uma fórmula sem quantificares na linguagem dos anéis ordenados e sejam $a_1, \ldots, a_n \in A$. Queremos mostrar que se existe $b \in K$ tal que $K \models \phi(a_1, \ldots, a_n, b)$, então existe $b' \in F$ tal que $F \models \phi(a_1, \ldots, a_n, b)$. Como existe um isomorfismo f de corpos ordenados entre F e o fecho real de $K \cap L$ em L, teremos que $L \models \phi(a_1, \ldots, a_n, f(b')$, o que concluirá a demonstração.

Pela *Regra de DeMorgan*, temos que

$$\phi(x_1, \ldots, x_n, y) \leftrightarrow \bigvee_{j=0}^{r} \bigwedge_{i=0}^{s} \theta_{i,j}(x_1, \ldots, x_n, y),$$

onde cada $\theta_{i,j}$ é uma fórmula atômica ou a negação de uma fórmula atômica. Como para satisfazer uma disjunção basta satisfazer uma das sentenças envolvidas, logo podemos supor, sem perda de generalidade, que $\phi(x_1, \ldots, x_n, y) \leftrightarrow \bigwedge_{i=0}^{s} \theta_i(x_1, \ldots, x_n, y)$.

Como os símbolos $\neq$ e $\leq$ podem ser escritos usando somente $=$ e $>$ nossa fórmula é equivalente a uma da forma:

$$\bigwedge_{i=1}^{t} (p_i(y) = 0) \wedge \bigwedge_{j=1}^{u} (q_j(y) > 0),$$

com $p_1, \ldots, p_t, q_1, \ldots, p_u \in A[x_1, \ldots, x_n, y]$. Se algum dos p_i for diferente de zero, então $b \in K$ é algébrico, logo $b' \doteq b \in F$.

Só nos resta o caso em que $\phi(x_1, \ldots, x_n, y)$ é equivalente a $\bigwedge_{j=1}^{u}(q_j(y) > 0)$. Temos três casos: ou b é maior que todo elemento de F, ou é menor ou conseguimos achar c, d em F tais que $c < b < d$. Denotemos $d = +\infty$ no primeiro caso e $c = -\infty$ no segundo.

Cada polinômio q_j troca de sinal em apenas um número finito de pontos, todos estes em F. Dessa forma conseguimos, seguindo a notação acima, para cada q_j um intervalo $(c_j, d_j) \ni b$ onde q_j é positivo. Chamando de C o maior dos c_j (ainda menor que b) e de D o menor dos d_j (ainda maior que b), temos que todos os q_j são positivos no intervalo $(C, D) \subseteq F$. Ora, mas estes são todos não vazios, então basta tomar $b' \in (C, D)$ que o nosso problema é resolvido. ■

Definição 4.2.10. Dizemos que uma teoria T é *completa* se para toda sentença ϕ temos que $T \models \phi$ ou $T \models \neg\phi$.

Corolário 4.2.11. *A Teoria dos Corpos Reais Fechados na linguagem dos anéis ordenados é completa.*

Demonstração. Como toda sentença ϕ é equivalente a uma ψ sem quantificadores, e afirmações sem quantificadores são preservadas por extensão e por subestruturas, logo temos:

$$K \models \psi \iff \mathbb{Q} \models \psi \iff L \models \psi,$$

onde K e L são dois corpos reais fechados. ■

Teoria dos Corpos Reais Fechados na linguagem dos anéis

Proposição 4.2.12. *A teoria dos corpos reais fechados na linguagem dos anéis (sem ordem) não possui eliminação de quantificadores.*

Demonstração. Vamos tomar o corpo $\mathbb{R}(x)$ e encontrar dois fechos reais para ele. Chamemos de K o fecho real de $R(x)$ com a ordem onde $0 < x < \frac{1}{n}$, $\forall n \in \mathbb{N}\backslash\{0\}$, e de L o fecho real de $R(x)$ com a ordem onde $-\frac{1}{n} < x < 0$, $\forall n \in \mathbb{N}\backslash\{0\}$.

Notemos que o corpo $R(x)$ é subcorpo de L e de K, mas não é subcorpo ordenado de nenhum dos dois, pois as possíveis ordens de K e L não coincidem sobre $R(x)$.

Tomemos a afirmação $\phi(y) \doteq (\exists z)(y = z)$. Temos que $x \in R(X)$ e que $K \models \phi(x)$, mas $L \models \neg\phi(x)$. Assim $\phi(x)$ não admite eliminação de quantificadores. ■

A prova acima poderia ser feita sem o uso do fecho real ou de corpos ordenados não arquimedianos, mas de forma que talvez causasse um estranhamento ao leitor. Ao invés de $\mathbf{R}[x]$, poderíamos usar $\mathbb{Q}[\sqrt{2}]$, ou melhor, $\frac{Q[x]}{\langle x^2-2\rangle}$. Podemos incluir esse corpo em R de duas formas diferents: associando x à raiz positiva de 2 ou à raiz negativa de 2, causando ordens diferentes em $\frac{Q[x]}{\langle x^2-2\rangle}$. O restante da prova seria igual.

Apesar de não admitir eliminação de quantificadores, essa teoria também é completa, uma vez que a relação de ordem $\leq$ da versão da linguagem dos anéis ordenados pode ser descrita nela! A demonstração consiste em simplesmente trocar todas as instância de $t(x_1, \ldots, x_n) \leq t'(x_1, \ldots, x_n)$ nas fórmulas da linguagem dos anéis ordenados por $(\exists y)(t(x_1, \ldots, x_n) + y^2 = t'(x_1, \ldots, x_n))$ e usar o corolário 4.2.11.

4.2.2 Corpos Algebricamente Fechados

Lema 4.2.13. *Se K e F são corpos algebricamente fechados, $F \subseteq K$, $\phi(x_1, \ldots, x_n, y)$ é livre de quantificadores, $a_1, \ldots, a_n \in F$, $K \models \phi(a_1, \ldots, a_n, b)$ para algum $b \in K$, então $F \models (\exists y)\phi(a_1, \ldots, a_n, y)$.*

Demonstração. Pela *Regra de DeMorgan*, temos que

$$\phi(x_1, \ldots, x_n, y) \leftrightarrow \bigvee_{j=0}^{r} \bigwedge_{i=0}^{s} \theta_{i,j}(x_1, \ldots, x_n, y),$$

onde cada $\theta_{i,j}$ é uma fórmula atômica ou a negação de uma fórmula atômica. Como para satisfazer uma disjunção basta satisfazer uma das sentenças envolvidas, logo podemos supor, sem perda de generalidade, que $\phi(x_1, \ldots, x_n, y) \leftrightarrow \bigwedge_{i=0}^{n} \theta_i(x_1, \ldots, x_n, y)$.

Uma fórmula atômica na linguagem dos anéis é claramente equivalente a uma da forma

$$p(x_1, \ldots, x_n, y) = 0,$$

pois são apenas igualdades de termos e os termos são polinômios em $\mathbb{Z}[x_1, \ldots, x_n, y]$. Se $p(x_1, \ldots, x_n, y) \in \mathbb{Z}[x_1, \ldots, x_n, y]$, podemos ver $p(a_1, \ldots, a_n, y)$ como um polinômio em $F[X]$.

Então temos que existem polinômios $p_1, \ldots, p_n, q_1, \ldots, q_n \in F[X]$ tais que $\phi(a_1, \ldots, a_n, y)$ é equivalente a

$$\bigwedge_{i=0}^{n} p_i(a_1, \ldots, a_n, y) = 0 \wedge \bigwedge_{i=0}^{m} q_i(a_1, \ldots, a_n, y) \neq 0.$$

Se pelo menos um p_i é não nulo, então b é algébrico sobre F, e logo $b \in F$, pois este é algebricamente fechado. Só resta o caso em que $\phi(a_1, \ldots, a_n, y)$ é equivalente a $\bigwedge_{i=0}^{m} q_i \neq 0$. Mas cada q_i tem um número finito de raízes, e todo corpo algebricamente fechado é infinto. Logo existe $c \in F$ tal que $F \models \phi(a_1, \ldots, a_n, c)$. ■

Teorema 4.2.14. *A* Teoria dos Corpos Algebricamente Fechados *admite eliminação de quantificadores.*

Demonstração. Vamos aplicar o corolário 4.2.8. Sejam K e F corpos algebricamente fechados, A uma sub-estrutura de ambos, $a_1, \ldots, a_n \in A$, e $\phi(x_1, \ldots, x_n, y)$ uma fórmula sem quantificadores. Queremos mostrar que, se existe $b \in K$ tal que $K \models \phi(a_1, \ldots, a_n, b)$, então existe $c \in F$ tal que $F \models \phi(a_1, \ldots, a_n, c)$.

Temos $A \subseteq K \cap F$ e a intersecção de corpos é um corpo. Seja L o fecho algébrico de $K \cap F$. Como L é isomorfo a uma subestrutura de K, pelo lema anterior, $L \models (\exists y)\phi(a_1, \ldots, a_n, y)$. Mas L também é isomorfo a uma subestrutura de F. Logo temos que $F \models (\exists y)\phi(a_1, \ldots, a_n, y)$, como queríamos provar. ■

Corolário 4.2.15. *A* Teoria dos Corpos Algebricamente Fechados de característica p *(p primo ou* 0*) é completa.*

Demonstração. Como toda sentença ϕ é equivalente a uma ψ sem quantificadores, e afirmações sem quantificadores são preservadas por extensão e por subestruturas, logo temos:

$$K \models \psi \iff \mathbb{F}_p \models \psi \iff L \models \psi,$$

onde K e L são dois corpos de característica p, e $\mathbb{F}_p$ é o corpo primo com esta característica ($\mathbb{Z}_p$ se p é primo, $\mathbb{Q}$ se $p = 0$). ■

Como uma aplicação interessante dos resultados acima obtemos

Teorema 4.2.16 (Nullstellensatz)**.** *Seja $J \triangleleft k[x_1, \ldots, x_n]$, onde k é algebricamente fechado. Então $I(V(J)) = \sqrt{J}$.*

Demonstração. Já vimos que a única inclusão difícil é $I(V(J)) \subseteq \sqrt{J}$. Suponhamos que ela não vale, isto é, que existe $f \in k[x_1, \ldots, x_n]$ tal que se $a_1, \ldots, a_l \in V(J)$, então $f(a_1, \ldots, a_l) = 0$ mas $\forall n \in \mathbb{N}$, $f^n \notin J$. Dentre todos os ideais $\tilde{J} \supseteq J$ que mantém a propriede de não conter nenhuma potência de f, tomemos um maximal. Claramente existe, pois caso contrário teríamos uma cadeia estritamente crescente infinita, o que a *CCA* não permite. Chamemos este de P. O nome é esse pois trata-se de um ideal primo:

Suponha que $ab \in P$, mas nem a nem b estão em P. Pela maximalidade de P, existem $m_1 \in \mathbb{N}$, $g_1 \in k[x_1, \ldots, x_n]$ e $p_1 \in P$ tais que $f^{m_1} = p_1 + g_1 a$. Logo, $bf^n \in P$. Repetindo o argumento, existem $m_2 \in \mathbb{N}$, $g_2 \in k[x_1, \ldots, x_n]$ e $p_2 \in P$ tais que $f^{m_2} = p_2 + g_2 b$. Logo $f^{m_1+m_2} = f^{m_1}p_2 + g_2 g f^{m_1} \in P$, uma contradição.

Como P é primo, $\frac{k[x_1,\ldots,x_n]}{P}$ é um domínio. Tomemos seu corpo de frações e depois o fecho algébrico deste, o qual chamaremos de L. L é uma extensão de k. Chamemos de $b_1, \ldots, b_n$ a imagem de $x_1, \ldots, x_n$ em L. Sejam $p_1, \ldots, p_r$ geradores de P. Temos assim que $p_i(b_1, \ldots, b_n) = 0$, mas $f(b_1, \ldots, b_n) \neq 0$.

Agora vem a Teoria dos Modelos: como a Teoria dos Corpos Algebricamente Fechados admite eliminação de quantificadores, e a afirmação de que existe uma raiz para os r polinômios p_i mas não pra f é obviamente descritível como uma fórmula de primeira ordem, ela também é satisfeita em k, o que contradiz a hipótese de que $f(a) = 0$ se $a \in V(J)$, pois sabemos da existencia de uma a tal que $p(a) = 0 \quad \forall p \in P \supseteq J$, mas $f(a) \neq 0$. ■

4.3 Uma prova do Teorema da Compacidade

Na demonstração do teorema 4.2.3, usamos um resultado importante e surpreendente conhecido como *Teorema da Compacidade*.

Se fossemos falar de *Teoria da Prova*, poderíamos falar que uma teoria é consistente se, e somente se, toda subteoria *finita* é consiste. Lá esse é um resultado bem tranquilo, pois uma teoria só é inconsistente se dela é possível provar uma contradição e toda prova utiliza apenas um número finito de afirmações.

Nosso objetivo aqui é diferente, mas parecido. Vamos mostrar que uma teoria tem modelo se, e somente se, toda subteoria finita tem modelo.

Este resultado não é nem um pouco óbvio. Poderíamos, por exemplo, introduzir na assinatura dos anéis ordenados um símbolo de constante "a" e incluir a seguinte lista de axiomas à teoria de $\mathbb{Z}$: $1 < a$, $2 < a$, $3 < a$, $\cdots$. Para cada subconjunto finito, temos que $\mathbb{R}$ é um modelo, bastando interpretar a como um número grande o suficiente. Mas então teremos que existe um modelo não arquimediano com a mesma teoria de primeira ordem de $\mathbb{Z}$! Em particular, a linguagem de primeira ordem não consegue capturar a informação de $\mathbb{Z}$ é arquimediano.

Uma possível prova seria pelo *Teorema da Completude*, que diz que se uma teoria é consistente então tem modelo. O leitor interessado pode consultar [3]. Não é do nosso interesse aqui falar de dedução formal, então decidimos seguir outro caminho. Vamos construir um modelo de uma teoria T a partir dos modelos das subteorias finitas. Para isso vamos usar o conceito de *ultraproduto*.

Um ultraproduto de $\mathcal{L}$-estruturas é uma outra $\mathcal{L}$-estrutura com a seguinte propriedade interessante: uma fórmula de 1ª ordem no ultraproduto é verdadeira se, e somente se, for verdadeira em uma porção "relevante" da $\mathcal{L}$-estruturas originais. O conceito de "relevante" depende de uma informação adicional na construção do ultraproduto: um ultrafiltro.

4.3.1 Filtros e Ultrafiltros

O conceito de *filtro* serve selecionar os subconjuntos "*suficientemente grandes*" ou *relevantes*. O nosso conceito de suficientemente grande vai exigir que a intersecção de dois conjuntos suficientemente grandes se interceptem em um conjunto suficientemente grande.

Definição 4.3.1. Seja I um conjunto. Chamamos de *filtro* sobre I a qualquer conjunto $\mathcal{F} \subset 2^I$, $\mathcal{F} \neq \emptyset$, tal que:

(i) $\emptyset \notin \mathcal{F}$; ($ii$) Se $X, Y \in \mathcal{F}$, então $X \cap Y \in \mathcal{F}$;

(iii) Se $X \in \mathcal{F}$ e $Y \in 2^I$, com $X \subset Y$, então $Y \in \mathcal{F}$.

Vamos dar dois exemplos de filtros para clarificar a ideia:

Exemplo 4.3.2. a) Seja I um conjunto *infinito*. É fácil verificar que o conjunto dos sub-conjuntos *cofinitos* em Y, $\mathcal{C} = \{X \in 2^I \mid I \backslash X \text{ é finito}\}$, é um filtro sobre I.

b) Sejam I um conjunto não vazio e $a \in I$ um elemento fixado. Então, o conjunto $\mathcal{P} = \{X \in 2^I \mid a \in X\}$ é um filtro sobre I, o filtro *principal* gerado por $\{a\}$.

No exemplo 4.3.2.(b), temos um peso especial para o elemento a. Ele tem um poder "ditatorial": um conjunto só é suficientemente grande se tiver a como elemento. Os filtros principais vão ser de pouca importância para nós, mas eles tem uma propriedade essencial que o filtro do exemplo 4.3.2 não tem.

Definição 4.3.3. Seja $\mathcal{U}$ um filtro sobre um conjunto I. Dizemos que $\mathcal{U}$ é um *filtro primo* se $A \cup B \in \mathcal{U}$, com $A, B \subset I$, então ou $A \in \mathcal{U}$ ou $B \in \mathcal{U}$.

Definição 4.3.4. Um filtro $\mathcal{U}$ sobre I é dito um *ultrafiltro* se ele é *maximal*, isto é, se dado algum filtro (sobre I) $\mathcal{F}$ tal que $\mathcal{U} \subseteq \mathcal{F}$, então $\mathcal{U} = \mathcal{F}$.

Proposição 4.3.5. *Seja I um conjunto. Para todo filtro $\mathcal{F}$ sobre I, existe um* ultrafiltro $\mathcal{U}$ *sobre I tal que $\mathcal{F} \subset \mathcal{U}$.*

Demonstração. Usaremos o *Lema de Zorn* para provarmos que o conjunto F de todos os filtros sobre I contendo $\mathcal{F}$ tem elemento maximal. F é certamente não vazio, pois $\mathcal{F} \in F$.

Seja $G \subset F$ um conjunto linearmente ordenado. Provaremos que $\mathcal{G} = \bigcup G$ é um filtro. Assim, como $\forall \mathcal{K} \in G$, $\mathcal{K} \subset \mathcal{G}$, o Lema de Kuratowsky-Zorn garante a existência de $\mathcal{U}$.

Com efeito, $\mathcal{G}$ é um filtro, pois:

(i) Se $\emptyset \in \mathcal{G}$ é porque está em algum $\mathcal{K} \in G$. Logo, $\emptyset \notin \mathcal{G}$;

(ii) Se $X, Y \in \mathcal{G}$, então cada um está em algum $\mathcal{K}_i \in G$, $i = 0, 1$. Chamemos de $\mathcal{K}$ o maior dos dois (existe pois o G é linearmente ordenado). Então $X, Y \in \mathcal{K}$, e como $X \cup Y \in \mathcal{K}$ temos que $X \cup Y \in \mathcal{G}$;

(iii) Se $X \in \mathcal{G}$, então $X \in \mathcal{K}$ para algum $\mathcal{K} \in G$. Logo, se $Y \in 2^I$ e $X \subset Y$, temos que $Y \in \mathcal{K}$ e, portanto, $Y \in \mathcal{G}$. ■

Assim existe pelo menos um ultrafiltro que contém o filtro do exemplo 4.3.2.(a). Vamos mostrar agora que todo ultrafiltro é um filtro primo.

Lema 4.3.6. *Seja $\mathcal{A} \subset 2^I$ uma família de conjuntos que satisfaz a* Propriedade da Intersecção Finita *(**PIF**), isto é, que a intersecção de qualquer quantidade finita de conjuntos-elementos de $\mathcal{A}$ é não-vazia. Então existe $\mathcal{F}$ um filtro sobre I tal que $\mathcal{A} \subset \mathcal{F}$.*

Demonstração. Seja $\tilde{\mathcal{A}}$ o conjunto de todas as intersecções finitas de elementos de $\mathcal{A}$. Definamos $\mathcal{F} \doteq \{B \in \wp(I) \ / \ \exists \tilde{B} \in \tilde{\mathcal{A}} \text{ tal que } \tilde{B} \subset B\}$.

Provemos que $\mathcal{F}$ é um filtro:

(i) $\emptyset \notin \mathcal{F}$ pois se fosse o caso teríamos que $\emptyset \in \tilde{\mathcal{A}}$, contrariando a PIF;

(ii) Se $X, Y \in \mathcal{F}$, existem $\tilde{X}, \tilde{Y} \in \tilde{\mathcal{A}}$ tais que $\tilde{X} \subset X$ e $\tilde{Y} \subset Y$. Então $\tilde{X} \cap \tilde{Y} \in \tilde{\mathcal{A}}$ e $\tilde{X} \cup \tilde{Y} \subset X \cap Y$;

(iii) Se $X \in \mathcal{F}$, então existe $\tilde{X} \in \tilde{\mathcal{A}}$. Seja $Y \in 2^I$, com $X \subset Y$. Então temos que $\tilde{X} \subset Y$ e, portanto, $Y \in \mathcal{F}$. ■

Lema 4.3.7. *Se $\mathcal{U}$ é um ultrafiltro sobre I, então também é um filtro primo.*

Demonstração. Suponha, para $A, B \subseteq I$, que $A \cup B$ está em $\mathcal{U}$, mas que tenhamos $A, B \notin \mathcal{U}$. Como $\mathcal{U}$ é maximal, os filtros gerados por $\mathcal{U} \cup \{A\}$ e $\mathcal{U} \cup \{B\}$ são ambos iguais a 2^I, isto é, pela prova do Lema 4.3.6, existem $X, Y \in \mathcal{U}$ tais que $X \cap A = \emptyset = Y \cap B$. Sem perda de generalidade, podemos tomar $X = Y$ (basta tomar a intersecção dos dois). Assim temos que $\emptyset = (X \cap A) \cup (X \cap B) = X \cap (A \cup B)$, o que contradiz $A \cup B \in \mathcal{U}$. ■

4.3.2 Ultraprodutos

Agora vamos usar essas informações em *Teoria dos Modelos*:

Definição 4.3.8. Seja $\mathcal{L}$ uma linguagem de primeira ordem com igualdade, I um conjunto, $\{\mathcal{M}_i\}_{i\in I}$ uma família indexada de $\mathcal{L}$-estruturas e $\mathcal{U}$ um ultrafiltro sobre I. Definimos em

$$\prod_{i\in I} \mathcal{M}_i$$

a seguinte relação, $\sim$:

$$(a_i)_{i\in I} \sim (b_i)_{i\in I} \iff \{i \in I \mid a_i = b_i\} \in \mathcal{U}$$

Segue facilmente do fato de que $\mathcal{U}$ é um **filtro** que $\sim$ é uma relação de equivalência. Denotaremos o quociente por

$$\prod_{\mathcal{U}} \mathcal{M}_i.$$

Notação 4.3.9. Vamos denotar a classe de equivalência de $(a_i)_{i\in I}$ em $\prod_{\mathcal{U}} \mathcal{M}_i$ por $[(a_i)_{i\in I}]$

O conjunto $\prod_{\mathcal{U}} \mathcal{M}_i$ torna-se uma $\mathcal{L}$-estrutura com as seguintes definições :

(i) Se $c \in \mathcal{L}$ é um símbolo de constante, sua interpretação será $[(c^{\mathcal{M}_i})_{i\in I}]$;

(ii) Se $f \in \mathcal{L}$ é um símbolo função n-ária, e $a = ([(a_{i1})_{i\in I}], \ldots, [(a_{in})_{i\in I}])$ é uma n-upla, a interpretação de $f(a)$ é $[(f(a_{i1}, \ldots, a_{in}))_{i\in I}]$;

(iii) Se $R \in \mathcal{L}$ é um símbolo de relação n-ária, $\prod_{\mathcal{U}} \mathcal{M}_i \models R([(a_{i1})_{i\in I}], \ldots, [(a_{in})_{i\in I}])$ se $\{i \in I \mid (a_{i1}, \ldots, a_{in}) \in R^{\mathcal{M}_i}\} \in \mathcal{U}$.

A $\mathcal{L}$-estrutura $\prod_{\mathcal{U}} \mathcal{M}_i$ denomina-se **ultraproduto** dos M_i módulo $\mathcal{U}$.

Para validar a definição, não é difícil mostrar que o item (ii) está bem definido: para cada coordenada $j \in \{1, \ldots, n\}$ temos que se $(a_{ij})_{i\in I} \sim (b_{ij})_{i\in I}$ então existe $S_j \in \mathcal{U}$ tal que $a_i = b_i$ se $i \in S_j$. Assim $S = \bigcap_{j=1}^{n} S_j \in \mathcal{U}$ e $f(a_{i1}, \ldots, a_{in}) = f(b_{i1}, \ldots, b_{in})$ se $i \in S$.

A importância da construção em 4.3.8 está no Teorema abaixo:

Teorema 4.3.10 (Teorema Fundamental dos Ultraprodutos, Łós)**.** *Se ϕ é uma sentença de $\mathcal{L}$, $\{M_i : i \in I\}$ é uma família de $\mathcal{L}$-estruturas e $\mathcal{U}$ é um ultrafiltro em I, então*

$$\prod_{\mathcal{U}} \mathcal{M}_i \models \phi \iff \{i \in I \mid \mathcal{M}_i \models \phi\} \in \mathcal{U}.$$

Demonstração. Provaremos algo ligeiramente mais geral. Sejam $a_i \in \mathcal{M}_i$ elementos quaisquer de $\mathcal{M}_i$ e $\phi(v)$ uma $\mathcal{L}$-fórmula (os a_i e v podem ser também sequencias finitas, mas escrevemos assim para não carregar a notação). Provaremos que

$$\prod_{\mathcal{U}} \mathcal{M}_i \models \phi([(a_i)_{i\in I}]) \iff \{i \in I \mid \mathcal{M}_i \models \phi(a_i)\} \in \mathcal{U}$$

Seguiremos por indução na complexidade. Se ϕ é atômica, o resultado é consequência imediata da definição de ultraproduto. Se já sabemos o resultado para duas fórmulas ϕ e ψ:

1. ($\wedge$)

$$\prod_{\mathcal{U}} \mathcal{M}_i \models \phi \wedge \psi([(a_i)_{i\in I}]) \iff \{i \in I \ / \ \mathcal{M}_i \models \phi(a_i)\} \in \mathcal{U} \text{ e } \{i \in I \ / \ \mathcal{M}_i \models \psi(a_i)\} \in \mathcal{U}$$

$$\iff \text{(usando a propriedade da intersecção dos filtros) } \{i \in I \ / \ \mathcal{M}_i \models \phi \wedge \psi(a_i)\} \in \mathcal{U}$$

2. ($\neg$) Aqui precisamos do *ultra*filtro:

$$\prod_{\mathcal{U}} \mathcal{M}_i \not\models \phi([(a_i)_{i\in I}]) \iff \{i \in I \mid \mathcal{M}_i \models \phi(a_i)\} \notin \mathcal{U}$$

$$\iff \text{(usando o Lema 4.3.7) } \{i \in I \ / \ \mathcal{M}_i \models \neg\phi(a_i)\} \in \mathcal{U}$$

3. ($\exists$) Se x é uma variável livre em ϕ e $\varphi(v) = \exists x \phi(v, x)$, então temos:

$$M_i \models \varphi(a_i), i \in A \in \mathcal{U} \iff \text{existe } b_i \text{ tal que } M_i \models \phi(a_i, b_i), i \in A \in \mathcal{U}$$

$$\iff \text{existe } (b_i)_{i\in I} \text{ tal que } \prod_{\mathcal{U}} \mathcal{M}_i \models \phi([(a_i)_{i\in I}], [(b_i)_{i\in I}]) \iff \prod_{\mathcal{U}} \mathcal{M}_i \models \varphi([(a_i)_{i\in I}])$$

Sabemos que todos os outros símbolos lógicos podem ser escritos em termos destes, completando a demonstração. ■

Com o teorema 4.3.10, é fácil demonstrar o Teorema da Compacidade:

Teorema 4.3.11 (Teorema da Compacidade)**.** *Seja* Γ *um conjunto de sentenças tal que todo subconjunto* finito *de* Γ *seja satisfazível. Então* Γ *é satisfazível.*

Demonstração. Seja I o conjunto dos subconjuntos finitos de Γ, e para cada $A \in I$, seja $\mathcal{M}_A$ um modelo de A. Faremos o ultraproduto dos M_A usando um ultrafiltro conveniente. Bastaria que esse ultrafiltro tivesse, para cada $\phi \in \Gamma$ o conjunto $C_\phi = \{A \in I \mid \phi \in A\}$. Pelas proposições 4.3.6 e 4.3.5, basta ver que o conjunto de todas C_ϕ satisfaz a *PIF*. Mas isto é simples, pois $\{\phi_1, \ldots, \phi_n\} \in C_{\phi_1} \cap \ldots \cap C_{\phi_n}$. ■

Referências Bibliográficas

[1] Saugata Basu, Richard Pollack, Marie-Françoise Roy, **Algorithms in Real Algebraic Geometry**, Algorithms and Computation in Mathematics, Volume 10, Springer-Verlag Berlin Heidelberg, 2003.

[2] Jacek Bochnak, Michel Coste, Marie-Francoise Roy, **Real Algebraic Geometry**, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3, Volume 36, Springer-Verlag Berlin Heidelberg, New York, 1998. 30

[3] C. C. Chang, H. J. Keisler, **Model Theory**, North-Holland Publ. Co., Amsterdam, 1990. 37, 44

[4] R. Engelking, **General Topology**, Sigma Series in Pure Math. **6**, Heldermann Verlag, Berlin, 1989. 8

[5] N. Jacobson, **Basic Algebra II**, 2nd edition, W. H. Freeman and Co., New York, 1989. 26

[6] S. Lang, **Algebra**, third edition, Springer-Verlag, New York, 2002. 26

[7] E. I. Korkina, A. G. Kushnirenko, *Another proof of the Tarski-Seidenberg theorem*, Siberian Mathematical Journal, **26** (1985), 703-707.

[8] M. Marshall, *Tarski-Seidenberg Theorem for Beginners*, math.usask.ca/ marshall/ts.ps.

[9] Victoria L. Noquez, **Model Theory of Real Closed Fields**, Logic and Computation Senior Thesis, Carnegie Mellon University, 2008.

[10] Hourya Sinaceur, **Corps Et Modeles: Essai Sur L'Histoire de L'Algebre Reelle (Mathesis)**, Vrin, Germany, 1991.

[11] Hourya Sinaceur, *Deux moments dans l'histoire du Théorème d'algèbre de Ch. F. Sturm*, Revue d'Histoire des Sciences., **48** (1998), 99-132.

[12] David Eugene Smith, Marcia L. Latham, **The geometry of René Descartes**, translated from the French and Latin, Dover Publications, New York, 1954.

Índice Remissivo