

O TEOREMA DE GAUSS-LUCAS

O problema da localização das raízes de um polinómio é um tema central no estudo de equações de variável real ou complexa. Esta localização torna-se gradualmente mais difícil à medida que o grau do polinómio aumenta, pelo que conseguir relacionar as raízes de um polinómio com as da sua derivada, cujo grau tem uma unidade a menos, pode ser bastante útil. Neste Canto Délfico vamos explorar o Teorema de Gauss-Lucas, um resultado clássico de Análise Complexa, que afirma que os zeros da derivada de um polinómio, quando representados no plano complexo, estão situados no invólucro convexo formado pelos zeros do polinómio.

1. INTRODUÇÃO

Neste Canto Délfico vamos analisar a relação entre os zeros de um polinómio complexo e os zeros da sua derivada. Esta relação pode ser vista como uma generalização aos polinómios complexos do Teorema de Rolle da Análise Real, provado por Michel Rolle no livro *Traité d'Algebre* em 1690, que afirma que entre dois zeros consecutivos de uma função diferenciável, existe pelo menos um zero da derivada:

Teorema 1.1 (Teorema de Rolle) *Se f é uma função real contínua no intervalo $[a, b]$ e diferenciável em (a, b) com $f(a) = f(b)$, então existe um número $a < c < b$ tal que $f'(c) = 0$.*

Se $p : \mathbb{R} \rightarrow \mathbb{R}$ for um polinómio de grau n com n raízes reais distintas, o Teorema de Rolle garante que a derivada p' tem exatamente $n - 1$ raízes reais. Além disso, as raízes de p e da sua derivada estão interlaçadas, ou seja, entre duas raízes consecutivas de p existe exatamente uma raiz de p' . Portanto, se x_1, \dots, x_n forem as raízes de p , os seus pontos críticos (i.e. os zeros de p') estão situados no segmento de reta formado pelos pontos $\min_{1 \leq i \leq n} x_i$ e $\max_{1 \leq i \leq n} x_i$.

Ao longo dos anos houve várias tentativas de generalizar este resultado para polinómios complexos [4]. No entanto, esta generalização não é óbvia nem imediata. Por exemplo, o polinómio

$$p(z) = (z^2 + 1)(z - 1)$$

tem como raízes os pontos $1 \pm i$, vértices de um triângulo isósceles. Se o Teorema de Rolle fosse válido no plano complexo, cada lado deste triângulo deveria ter um ponto crítico de p , o que não acontece pois p tem apenas dois pontos críticos no interior do triângulo. No plano complexo, o conceito de ponto crítico situado entre duas raízes de um polinómio é geralmente substituído pelo conceito mais geral de uma região convexa definida pelas raízes do polinómio onde se localizam os pontos críticos.

Neste sentido, o Teorema de Gauss-Lucas generaliza o Teorema de Rolle para polinómios complexos, afirmando que os pontos críticos de p , quando considerados como pontos no plano complexo, estão situados no interior do mais pequeno conjunto convexo que contém os zeros de p . É um resultado clássico da Análise Complexa, com várias generalizações e demonstrações, que permite restrin-

gir a localização dos pontos críticos de um polinómio sem termos de calcular a sua derivada.

O Teorema de Gauss-Lucas deve o seu nome a Johann Carl Friedrich Gauss e a Felix Lucas. Gauss, entre 1836 e 1846, escreveu sobre este teorema ao estabelecer uma interpretação mecânica para os pontos críticos de um polinómio, mas foi Lucas em 1870, aparentemente sem conhecer o resultado de Gauss, quem apresentou a primeira prova do teorema [5].

2. POLINÓMIOS COMPLEXOS

No que se segue, assumimos uma certa familiaridade com os números complexos. Identificamos um número complexo $z = a + bi$, onde a e b são números reais e $i^2 = -1$, com o ponto (a, b) do plano \mathbb{R}^2 . O conjugado de $z = a + bi$ é o número $\bar{z} = a - bi$, obtido de z por reflexão sobre o eixo real (veja-se a Figura 1). A distância de $z = a + bi$ à origem é dada pelo número $|z| = \sqrt{a^2 + b^2}$. É fácil verificar que $z\bar{z} = |z|^2$.

Um polinómio complexo de grau $n \geq 1$ é uma função $p : \mathbb{C} \rightarrow \mathbb{C}$ da forma

$$p(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0, \quad (1)$$

com $a_0, a_1, \dots, a_n \in \mathbb{C}$ e $a_n \neq 0$. Um número complexo α que verifique $p(\alpha) = 0$ chama-se raiz do polinómio. O Teorema Fundamental da Álgebra garante que um polinómio de grau $n \geq 1$ tem exatamente n raízes complexas, contando multiplicidades. Por exemplo,

$$p(z) = (z^2 + 1)(z - 1) \quad (2)$$

é um polinómio de grau 3 e tem raízes complexas dadas por $z_1 = -i$, $z_2 = i$ e $z_3 = 1$.

Um polinómio complexo é uma função diferenciável em todo o plano complexo. A derivada de (1) é o polinómio de grau $n - 1$

$$p'(z) = n a_n z^{n-1} + (n-1) a_{n-1} z^{n-2} + \dots + a_1.$$

As raízes de $p'(z)$ dizem-se pontos críticos de $p(z)$. Por exemplo, a derivada do polinómio (2) é o polinómio

$$p'(z) = 3z^2 - 2z + 1.$$

Usando a Fórmula Resolvente, obtemos as raízes $\zeta_1 = \frac{1}{3} + i\frac{\sqrt{2}}{3}$ e $\zeta_2 = \frac{1}{3} - i\frac{\sqrt{2}}{3}$ de $p'(z)$. Não é difícil verificar que os pontos críticos $\frac{1}{3} \pm i\frac{\sqrt{2}}{3}$ do polinómio (2) estão no interior do triângulo cujos vértices são as raízes do polinómio, como se pode confirmar na Figura 2, onde as raízes do polinómio estão representadas a azul e os pontos críticos a cinzento.

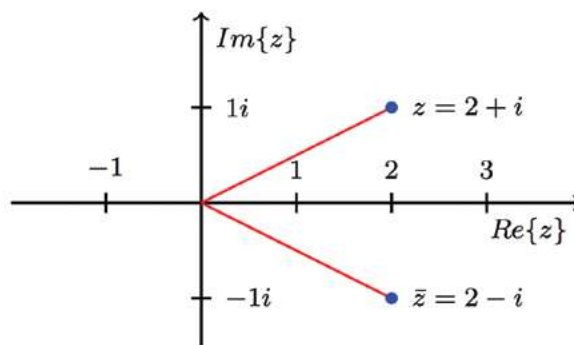


Figura 1. Representação do número $2 + i$ e do seu conjugado no plano complexo.

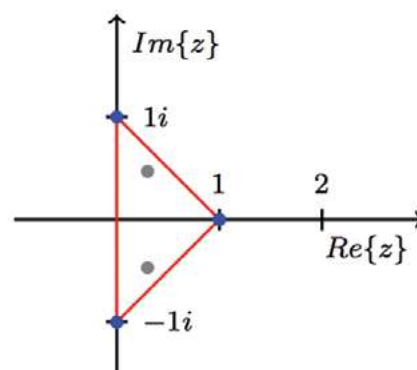


Figura 2. Raízes e pontos críticos do polinómio $p(z) = (z^2 + 1)(z - 1)$.

3. CONJUNTOS CONVEXOS

Nesta secção vamos introduzir o conceito de conjunto convexo e de envólucro convexo de um conjunto de números (veja-se, por exemplo [7]). Apesar da definição e dos resultados que vamos apresentar serem válidos em geral, vamos restringir-nos a conjuntos de números complexos.

Intuitivamente, dizemos que um conjunto $S \subset \mathbb{C}$ é convexo se contém o segmento de reta entre dois quaisquer pontos de S . O triângulo S_2 representado na Figura 3 é um exemplo de conjunto convexo, enquanto que o conjunto S_1 não é convexo, uma vez que o segmento de reta que une z a w não está contido em S_1 .

A reta que passa pelos números complexos z e w , pode ser dada parametricamente por

$$r(\lambda) = z + \lambda(w - z) = (1 - \lambda)z + \lambda w,$$

com $\lambda \in \mathbb{R}$. Notemos que quando $\lambda = 0$ obtemos o ponto $r(0) = z$, e quando $\lambda = 1$ obtemos o ponto $r(1) = w$. Assim, o segmento de reta que une z a w , denotado por $[z, w]$,

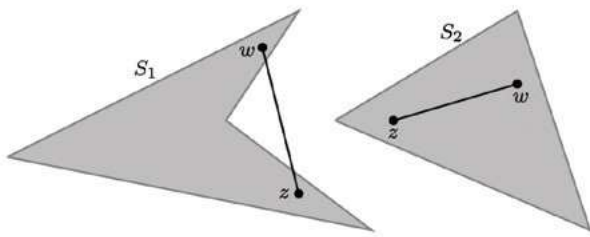


Figura 3. Um conjunto não convexo e um conjunto convexo.

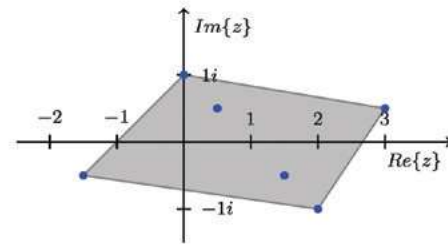


Figura 4. O invólucro convexo das raízes do polinómio (3).

pode ser representado por

$$[z, w] = \{(1 - \lambda)z + \lambda w : 0 \leq \lambda \leq 1\}.$$

Ou seja, os números pertencentes ao segmento $[z, w]$ podem ser escritos na forma $\lambda_1 z + \lambda_2 w$, com $\lambda_i \geq 0$ e $\lambda_1 + \lambda_2 = 1$. Um tal número diz-se uma combinação convexa de z e w . Mais geralmente, uma combinação convexa de n pontos z_1, \dots, z_n é uma soma da forma

$$\lambda_1 z_1 + \dots + \lambda_n z_n,$$

onda cada $\lambda_i \geq 0$ e $\lambda_1 + \dots + \lambda_n = 1$.

Podemos agora apresentar a definição formal de conjunto convexo.

Definição 1. Um conjunto $S \subset \mathbb{C}$ é convexo se $[z, w] \subset S$ para quaisquer $z, w \in S$.

É fácil verificar que a união de conjuntos convexas não é, em geral, convexo. Por exemplo, a união dos segmentos $[0, 1]$ e $[0, i]$ não é convexo. No entanto, a interseção de conjuntos convexas é ainda convexo.

Proposição 3.1. A interseção de um qualquer número de conjuntos convexas é um conjunto convexo.

Demonstração. Sejam $\{S_i : i \in I\}$ uma família de conjuntos convexas e $S = \bigcap_{i \in I} S_i$. Então, dados $z, w \in S$ temos $z, w \in S_i$ para todo o $i \in I$. Como cada S_i é convexo, verifica-se $[z, w] \subset S_i$, pelo que $[z, w] \subset S$. Ou seja, S é convexo. \square

Intuitivamente, a interseção de todos os conjuntos convexas que contêm S é o "menor" conjunto convexo que contém S . Esta é a definição de invólucro convexo de S .

Definição 2. O invólucro convexo de um conjunto S , denotado por $\text{conv}(S)$, é a interseção de todos os conjuntos convexas que contêm S .

Podemos pensar no invólucro convexo de um conjun-

to de pontos $z_1, \dots, z_n \in \mathbb{C}$, como sendo o conjunto que se obtém colocando pinos em cada um dos pontos z_i e rodeando todos eles com um elástico esticado. A área demarcada corresponde ao invólucro convexo $\text{conv}(z_1, \dots, z_n)$. Por exemplo, o invólucro convexo do conjunto formado pelas raízes do polinómio de grau 6

$$p(z) = \prod_{i=1}^6 (z - z_i), \quad (3)$$

com $z_1 = -3/2 - i/2$, $z_2 = i$, $z_3 = 3 + i/2$, $z_4 = 2 - i$, $z_5 = 1/2 + i/2$ e $z_6 = 3/2 - i/2$ pode ser visualizado na Figura 4.

Por definição, um conjunto convexo contém todas as combinações convexas de dois dos seus pontos. Na próxima proposição, vamos mostrar que este resultado se generaliza para combinações convexas de mais do que dois pontos.

Proposição 3.2. Se S é um conjunto convexo, então qualquer combinação convexa (finita) de elementos de S pertence a S .

Demonstração. Vamos efetuar a prova por indução sobre o número n de pontos da combinação convexa. Os casos $n = 1$ e $n = 2$ resultam da definição de conjunto convexo. Suponhamos a proposição verdadeira para combinações convexas com $n = k$ pontos de S e consideremos a combinação convexa

$$\lambda_1 z_1 + \dots + \lambda_k z_k + \lambda_{k+1} z_{k+1}. \quad (4)$$

Seja $\lambda = \sum_{i=1}^k \lambda_i$. Então

$$1 - \lambda = \sum_{i=1}^{k+1} \lambda_i - \sum_{i=1}^k \lambda_i = \lambda_{k+1}$$

e podemos escrever

$$(4) = \lambda \left(\sum_{i=1}^k \frac{\lambda_i}{\lambda} z_i \right) + \lambda_{k+1} z_{k+1}. \quad (5)$$

Uma vez que $\sum_{i=1}^k \frac{\lambda_i}{\lambda} = 1$, a hipótese indutiva garante que a combinação convexa $\sum_{i=1}^k \frac{\lambda_i}{\lambda} z_i$ pertence ao conjunto S . Mas então, o lado esquerdo da igualdade (5) é uma combinação convexa de dois pontos de S e, portanto, pertence a S . \square

Podemos usar combinações convexas para descrever o invólucro convexo de um conjunto.

Teorema 3.3 *Seja $S \subset \mathbb{C}$. Então, o invólucro convexo de S é formado por todas as combinações convexas (finitas) de pontos de S .*

Demonstração. Seja $C(S)$ o conjunto formado por todas as combinações convexas (finitas) de pontos de S . Se $z \in C(S)$, então z é uma combinação convexa de pontos de S . Como $S \subset \text{conv}(S)$ e $\text{conv}(S)$ é convexo, segue da Proposição 3.2 que $z \in \text{conv}(S)$. Portanto, temos $C(S) \subset \text{conv}(S)$. Para provarmos a inclusão contrária basta mostrarmos que é um conjunto convexo, pois neste caso pela definição de invólucro convexo temos $\text{conv}(S) \subset C(S)$.

Sejam então $z, w \in C(S)$. Isto significa que

$$z = \sum_{i=1}^n \lambda_i z_i \quad \text{e} \quad w = \sum_{i=1}^m \mu_i w_i,$$

para certos pontos $z_1, \dots, z_n, w_1, \dots, w_m \in S$ com $\sum_{i=1}^n \lambda_i = \sum_{i=1}^m \mu_i = 1$. Seja $\lambda \in [0, 1]$ e consideremos um ponto do segmento $[z, w]$:

$$(1 - \lambda)z + \lambda w. \quad (6)$$

Usando as representações de z e w , temos

$$\begin{aligned} (6) &= (1 - \lambda) \left(\sum_{i=1}^n \lambda_i z_i \right) + \lambda \left(\sum_{i=1}^m \mu_i w_i \right) \\ &= \sum_{i=1}^n (1 - \lambda) \lambda_i z_i + \sum_{i=1}^m \lambda \mu_i w_i, \end{aligned}$$

com

$$\sum_{i=1}^n (1 - \lambda) \lambda_i + \sum_{i=1}^m \lambda \mu_i = (1 - \lambda) + \lambda = 1.$$

Ou seja, o ponto $(1 - \lambda)z + \lambda w$ é uma combinação convexa de pontos de S , pelo que pertence a $C(S)$. Isto prova que o segmento $[z, w]$ está contido em $C(S)$ e, portanto, $C(S)$ é convexo. \square

4. O TEOREMA DE GAUSS-LUCAS

Estamos agora em condições de enunciar e provar o Te-

orema de Gauss-Lucas. A demonstração deste resultado pode ser obtida usando, por exemplo, argumentos de natureza analítica, geométrica ou física. Neste artigo vamos apresentar uma prova analítica do teorema e remetemos o leitor para as referências [2, 6, 8], onde poderão encontrar provas de natureza distintas.

Teorema 4.1 (Teorema de Gauss-Lucas). *Os pontos críticos de um polinômio não constante p pertencem ao invólucro convexo das raízes de p .*

Antes de apresentarmos a demonstração deste resultado, vamos analisar alguns casos particulares. O teorema é trivial se p tem grau 1, pois neste caso o polinômio não tem pontos críticos. Para polinômios de grau 2 com raízes distintas z_1 e z_2 , temos

$$\begin{aligned} p(z) &= a(z - z_1)(z - z_2) \\ &= a \left(z^2 - z(z_1 + z_2) + z_1 z_2 \right). \end{aligned}$$

para alguma constante não nula a . A partir daqui é fácil obter a derivada

$$p'(z) = a(2z - (z_1 + z_2))$$

e a sua única raiz:

$$\zeta = \frac{z_1 + z_2}{2}.$$

O invólucro convexo de z_1 e z_2 é o segmento $[z_1, z_2]$ e o ponto crítico é o ponto médio deste segmento.

Se p for um polinômio de grau 2 com apenas uma raiz z_1 de multiplicidade 2, então

$$p(z) = a(z - z_1)^2 \text{ e } p'(z) = 2a(z - z_1),$$

com $a \neq 0$. Neste caso, o ponto crítico de p coincide com a única raiz de p e $\text{conv}(z_1) = \{z_1\}$.

Demonstração do Teorema de Gauss-Lucas. Seja

$$p(z) = a \prod_{i=1}^n (z - z_i)$$

um polinômio complexo de grau $n \geq 1$, onde os números z_1, \dots, z_n são as suas raízes, não necessariamente distintas. Seja z_0 um ponto crítico de p . Se z_0 coincide com uma das raízes de p , é claro que está no invólucro convexo das raízes de p . Suponhamos então que $p(z_0) \neq 0$. Observe-mos que as raízes da função racional

$$\frac{p'(z)}{p(z)} = \sum_{i=1}^n \frac{1}{z - z_i},$$

são os pontos críticos de p . Em particular, temos

$$0 = \frac{p'(z_0)}{p(z_0)} = \sum_{i=1}^n \frac{1}{z_0 - z_i}.$$

Multiplicando e dividindo cada termo do membro direito da igualdade anterior por $\overline{z_0 - z_i}$ obtemos

$$0 = \sum_{i=1}^n \frac{\overline{z_0 - z_i}}{(z_0 - z_i)(\overline{z_0 - z_i})} = \sum_{i=1}^n \frac{\overline{z_0 - z_i}}{|z_0 - z_i|^2},$$

ou seja,

$$\sum_{i=1}^n \frac{\overline{z_0}}{|z_0 - z_i|^2} = \sum_{i=1}^n \frac{\overline{z_i}}{|z_0 - z_i|^2}.$$

Conjugando ambos os membros desta igualdade obtemos

$$\sum_{i=1}^n \frac{z_0}{|z_0 - z_i|^2} = \sum_{i=1}^n \frac{z_i}{|z_0 - z_i|^2},$$

ou ainda

$$z_0 \sum_{i=1}^n \frac{1}{|z_0 - z_i|^2} = \sum_{i=1}^n \frac{1}{|z_0 - z_i|^2} z_i.$$

Fazendo $\lambda = \sum_{i=1}^n \frac{1}{|z_0 - z_i|^2}$ conseguimos escrever z_0 como uma combinação convexa das raízes de p ,

$$z_0 = \sum_{i=1}^n \frac{\mu_i}{\lambda} z_i,$$

onde $\mu_i = \frac{1}{|z_0 - z_i|^2}$ para $1 \leq i \leq n$, uma vez que $\sum_{i=1}^n \frac{\mu_i}{\lambda} = 1$. Usando o Teorema 3.3 concluímos então que o ponto crítico z_0 pertence ao invólucro convexo das raízes de p . \square

Vamos ilustrar o Teorema de Gauss-Lucas analisando o polinómio

$$p(z) = \frac{1}{3}(z^2 - 6)(z - 3i). \quad (7)$$

O invólucro convexo das raízes $3i$ e $\pm\sqrt{6}$ é o triângulo cujos vértices são estas raízes. A derivada $p'(z) = z^2 - 2iz - 2$ tem raízes $\pm 1 + i$, que se encontram no interior do triângulo.

No caso de polinómios de grau 3, podemos indicar com mais exatidão a localização dos pontos críticos usando um resultado notável conhecido como Teorema de Steiner-Marden [1, 3].

Teorema 4.2 (Teorema de Steiner-Marden) *Se p é um po-*

linómio de grau 3 cujas raízes são não colineares no plano complexo, então os pontos críticos de p são os focos da elipse inscrita no triângulo formado pelas raízes de p e tangente aos lados deste triângulo nos seus pontos médios.

A elipse considerada no Teorema de Steiner-Marden é conhecida como elipse de Steiner. Uma prova da sua existência e unicidade pode ser encontrada em [3]. A Figura 5 representa a configuração geométrica desta elipse com focos nos pontos e e f , inscrita no triângulo de vértices a, b, c .

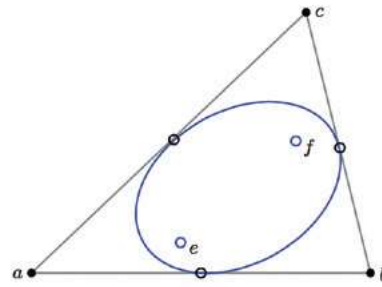


Figura 5. A elipse de Steiner.

Demonstração do Teorema 4.2. Suponhamos que a, b, c são as raízes do polinómio p e e, f são os seus pontos críticos. Então, podemos escrever

$$\begin{aligned} p(z) &= (z - a)(z - b)(z - c) \\ &= z^3 - (a + b + c)z^2 + \dots \end{aligned}$$

e

$$p'(z) = 3(z - e)(z - f) = 3(z^2 - (e + f)z + ef).$$

O Teorema de Gauss-Lucas garante que os pontos críticos pertencem ao interior do triângulo de vértices a, b, c .

Notemos que podemos alterar o sistema de coordenadas $z \mapsto z'$ aplicando a rotação $z \mapsto \alpha z$ com $|\alpha| = 1$, seguida da translação $z \mapsto z + \beta$. Nas novas coordenadas, o polinómio $p(z)$ transforma-se em $p(z') = p(\alpha z + \beta) =: q(z)$. Uma vez que $dq/dz = \alpha dp/dz$, os pontos críticos de p e de q coincidem e, portanto, podemos escolher um sistema de coordenadas onde o ponto 0 seja o baricentro do triângulo. Neste caso, temos

$$a + b + c = 0 \quad (8)$$

e, portanto, $e + f = 0$. A derivada de p pode então ser escrita como

$$p'(z) = (z - a)(z - b) + (2z - (a + b))(z - c) \quad (9)$$

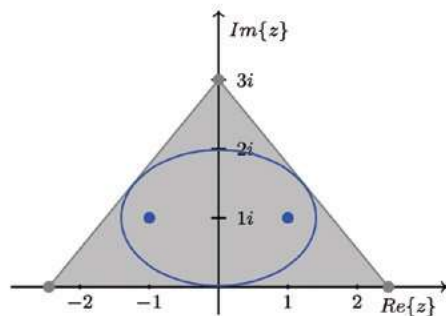


Figura 6. Raízes, pontos críticos e elipse de Steiner do polinómio $p(z) = \frac{1}{3}(z^2 - 6)(z - 3i)$.

mas também como

$$p'(z) = 3(z - e)(z + e). \quad (10)$$

Consideremos o ponto médio $z_1 = (a + b)/2$ do lado $[a, b]$ do triângulo. Usando as fórmulas (9) e (10) obtemos

$$3(z_1 - e)(z_1 + e) = -\left(\frac{a - b}{2}\right)^2. \quad (11)$$

Usando a identidade do paralelogramo e as fórmulas (8) e (11), obtemos a seguinte expressão para a soma das distâncias de z_1 aos pontos $-e$ e e :

$$\begin{aligned} 2(|z_1 + e| + |z_1 - e|)^2 &= \\ &= 2|z_1 + e|^2 + 2|z_1 - e|^2 + 4|(z_1 + e)(z_1 - e)| \\ &= 4|z_1|^2 + 4|e|^2 + \frac{1}{3}|a - b|^2 \\ &= |a + b|^2 + 4|e|^2 + \frac{1}{3}|a - b|^2 \\ &= \frac{1}{3}(|a + b|^2 + |a - b|^2) + \frac{2}{3}|a + b|^2 + 4|e|^2 \\ &= \frac{2}{3}(|a|^2 + |b|^2 + |c|^2) + 4|e|^2. \end{aligned}$$

Como esta última expressão é independente do lado do triângulo onde escolhemos o ponto médio z_1 , a elipse com focos $\pm e$, centro 0, e que passa pelo ponto médio $(a + b)/2$ também passa pelos pontos médios $(a + c)/2$ e $(b + c)/2$ dos outros lados do triângulo e, portanto, é a elipse de Steiner do triângulo de vértices a, b, c . \square

Voltando ao polinómio (7) considerado acima, é fácil verificar que o centro da elipse com focos $\pm 1 + i$, inscrita no triângulo de vértices $3i$ e $\pm\sqrt{6}$, tem centro i (baricentro do triângulo), menor eixo de comprimento 1 e maior eixo de comprimento $\sqrt{2}$. Portanto, a elipse de Steiner tem equação

$$\frac{x^2}{2} + (y - 1)^2 = 1,$$

como se pode ver na Figura 6.

AGRADECIMENTO

Agradeço ao Professor Alfredo Costa a leitura cuidadosa que fez a este documento.

REFERÊNCIAS

- [1] E. Badertscher (2014). A Simple Direct Proof of Marden's Theorem, *The American Mathematical Monthly*, 121:6, 547-548.
- [2] M. Brilleslyper and B. Schaubroeck (2017). Explorations of the Gauss-Lucas Theorem, *PRIMUS*, 27:8-9, 766-777.
- [3] D. Kalman (2008). An elementary proof of Marden's theorem, *Amer. Math. Monthly* 115, 330-338.
- [4] M. Marden (1985). The Search for a Rolle's Theorem in the Complex Domain, *The American Mathematical Monthly*, Vol. 92 (9), pp. 643-650.
- [5] M. Marden (1996). *Geometry of Polynomials*. Mathematical Surveys and Monographs, no. 3, American Mathematical Society.
- [6] A. Chéritat, Y. Gao, Y. Ou and L. Tan (2015). A refinement of the Gauss-Lucas theorem (after W. P. Thurston), *C. R. Math. Acad. Sci. Paris* 353, no. 8, 711-715.
- [7] V. Soltan (2020). *Lectures on Convex Sets*. Second Edition, World Scientific, Hackensack, NJ, 612 pp.
- [8] E. Van Vleck (1929). On the location of roots of polynomials and entire functions, *Bull. Amer. Math. Soc.* 35, 643-683.