A Note on the Eneström-Kakeya Theorem

Abstract. In this note, we give two proofs for a variant of the Eneström-Kakeya Theorem. Then we prove that this variant is equivalent to the strict Eneström-Kakeya Theorem as stated in Theorem 1.

1. Introduction

On March 9, 2019, the first author addressed a query posted by Seiichi Kirikami on a Yahoo Geometry Discussion Group (the archives of which are no longer extant due to platform changes). Although Kirikami's post was directed toward a member named 'Ed,' it was confirmed by the first author, then participating under the username 'ctperng.' Kirikami's query conjectured the validity of Corollary 1; the first author confirmed this using a proof fundamentally similar to the one presented for Theorem 2 herein. As it happens, Theorem 2 provides a natural generalization of Corollary 1 and is equivalent to Theorem 1.

The Eneström-Kakeya Theorem provides a fascinating example of how language and geography can delay the recognition of mathematical priority. The theorem was first established by the Swedish mathematician Gustaf Eneström [1] in 1893. Writing for a Swedish journal, Eneström derived the result while investigating the mathematical foundations of pension funds. However, because the paper was written in Swedisha language not widely read by the international mathematical community at the timehis work remained largely obscure for nearly two decades.

In 1912, the Japanese mathematician Soichi Kakeya [2] independently discovered the same result. Unlike Eneström, Kakeya published his proof in English in the Tôhoku Mathematical Journal. Given the accessibility of the language and the journals reach, the result quickly gained international traction and became known as "Kakeya's Theorem."

The hyphenated "Eneström-Kakeya" designation was adopted later to reconcile this history. Recognizing the chronological priority of the Swedish work, a colleague (T. Hayashi) urged Eneström [6] to publish a French translation of his original proof, which appeared in 1920. Today, the dual name honors both Eneström's chronological priority and Kakeya's role in broadening the reach of the theorem to the global stage.

2. Main Results

The primary results of this note are Theorems 2 and 3. We first assume Theorem 1 – a strict version of the Eneström-Kakeya Theorem – to provide an initial proof of Theorem 2 as a direct consequence. While Theorem 1 is a classical result that can be derived from the foundational works of Eneström [1] and Kakeya [2] (cf. [3], [4], and [5]), its external assumption is not required for our development.

Our second proof of Theorem 2 is established independently of Theorem 1, using Rouché's Theorem from complex analysis.

Since Theorem 3 subsequently establishes that Theorem 2 implies Theorem 1, the latter is formally recovered through our internal results. Consequently, the logical framework of this note is entirely self-contained.

Theorem 1. (The strict Eneström-Kakeya's Theorem) If $a_n > \cdots > a_1 > a_0 > 0$, the equation $a_n x^n + \cdots + a_1 x + a_0 = 0$ has all roots of absolute value < 1.

We give two proofs of Theorem 2 below. The first proof shows that it is a direct consequence of Theorem 1.

Theorem 2. If $a_{n-1} \geq \cdots \geq a_1 \geq a_0 > 0$ with $n > 1$ and $\sum_{k=0}^{n-1} a_k > 1$, then the equation

$$p(x) := x^n - (a_{n-1}x^{n-1} + \cdots + a_1 x + a_0) = 0$$

has a real root > 1 and $n-1$ roots of absolute value < 1.

First Proof of Theorem 2 Using Theorem 1. Since $p(1) < 0$ and $p(\infty) > 0$, it follows from the Intermediate Value Theorem and Descartes' rule of sign that the equation has a unique positive root $r > 1$, so one has the following identity

$$r^n - a_{n-1}r^{n-1} - \cdots - a_1 r - a_0 = 0. \tag{1}$$

By division, one may write

$$p(x) = (x - r)(b_{n-1}x^{n-1} + b_{n-2}x^{n-2} + \cdots + b_1 x + b_0), \tag{2}$$

where b_i's can be determined inductively as follows (using the relations (1) and (2)):

$$b_{n-1} = 1 = \frac{1}{r^n}[a_{n-1}r^{n-1} + \cdots + a_1 r + a_0]$$

$$b_{n-2} = r - a_{n-1} = \frac{1}{r^{n-1}}[r^n - a_{n-1}r^{n-1}] = \frac{1}{r^{n-1}}[a_{n-2}r^{n-2} + \cdots + a_1 r + a_0]$$

$$b_{n-3} = rb_{n-2} - a_{n-2} = r^2 - a_{n-1}r - a_{n-2} = \frac{1}{r^{n-2}}[a_{n-3}r^{n-3} + \cdots + a_1 r + a_0]$$

$$\cdots$$

$$b_1 = rb_2 - a_2 = r^{n-2} - a_{n-1}r^{n-3} - \cdots - a_2 = \frac{1}{r^2}[a_1 r + a_0]$$

$$b_0 = rb_1 - a_1 = r^{n-1} - a_{n-1}r^{n-2} - \cdots - a_1 = \frac{1}{r} \cdot a_0.$$

It follows from the above that

$$\frac{b_{n-j}}{b_{n-j-1}} = \frac{\frac{1}{r^{n-j+1}}[a_{n-j}r^{n-j} + a_{n-j-1}r^{n-j-1} + \cdots + a_1 r + a_0]}{\frac{1}{r^{n-j}}[a_{n-j-1}r^{n-j-1} + \cdots + a_1 r + a_0]}$$

$$= \frac{a_{n-j}r^{n-j-1}}{a_{n-j-1}r^{n-j-1} + \cdots + a_1 r + 1} + \frac{1}{r}, \text{ for } 1 \leq j \leq n-2$$

and

$$\frac{b_1}{b_0} = \frac{\frac{1}{r^2}[a_1 r + a_0]}{\frac{1}{r} \cdot a_0} = \frac{a_1}{a_0} + \frac{1}{r} > 1,$$

which fulfills the required condition for $n = 2$. Now we may assume that $n \geq 3$ and it remains to show that

$$\frac{b_{n-j}}{b_{n-j-1}} > 1 \text{ for } 1 \leq j \leq n-2.$$

But one has in this case

$$\frac{b_{n-j}}{b_{n-j-1}} = \frac{a_{n-j}r^{n-j-1}}{a_{n-j-1}r^{n-j-1} + \cdots + a_1 r + a_0} + \frac{1}{r} \geq \frac{a_{n-j}r^{n-j-1}}{a_{n-j-1}(r^{n-j-1} + \cdots + r + 1)} + \frac{1}{r}$$

$$= \frac{a_{n-j}r^{n-j-1}(r-1)}{a_{n-j-1}(r^{n-j} - 1)} + \frac{1}{r} \geq \frac{r^{n-j-1}(r-1)}{(r^{n-j} - 1)} + \frac{1}{r} = \frac{r^{n-j+1} - 1}{r^{n-j+1} - r} > 1,$$

as required. By Theorem 1, the theorem is proven. □

Second Proof of Theorem 2 Using Rouché's Theorem. As in the first proof, $p(z)$ has a unique real root $r > 1$. Consider the polynomial

$$(z-1)p(z) = f(z) + g(z),$$

where

$$f(z) = -(a_{n-1}+1)z^n, \qquad g(z) = z^{n+1} + \sum_{k=1}^{n-1}(a_k - a_{k-1})z^k + a_0.$$

On the unit circle $|z| = 1$, we have $|g(z)| \leq |f(z)|$ and equality holds only for $z = 1$, noting that $g(1) = 1 + a_{n-1}$ and using the triangle inequality. Consider the function $\psi(z) = |f(z)| - |g(z)|$. We know that $\psi(1) = 0$, and $\psi(z) > 0$ for all other z on the unit circle. Let

$$\varphi(r) = (a_{n-1}+1)r^n - \left[r^{n+1} + \sum_{k=1}^{n-1}(a_k - a_{k-1})r^k + a_0\right].$$

It is easy to see that

$$\varphi(1) = 0 \text{ and } \varphi'(1) = -1 + \sum_{k=0}^{n-1} a_k > 0.$$

This shows that there exists $\epsilon > 0$ such that $\varphi(1+\epsilon) > 0$, namely

$$(a_{n-1}+1)(1+\epsilon)^n - \left[(1+\epsilon)^{n+1} + \sum_{k=1}^{n-1}(a_k - a_{k-1})(1+\epsilon)^k + a_0\right] > 0.$$

Now on the circle $|z| = 1+\epsilon$, we have

$$|f(z)| = (a_{n-1}+1)(1+\epsilon)^n, \ |g(z)| \leq (1+\epsilon)^{n+1} + \sum_{k=1}^{n-1}(a_k - a_{k-1})(1+\epsilon)^k + a_0.$$

It follows that for $|z| = 1+\epsilon$, we have that

$$\begin{aligned}\psi(z) &= |f(z)| - |g(z)| \\ &\geq (a_{n-1}+1)(1+\epsilon)^n - \left[(1+\epsilon)^{n+1} + \sum_{k=1}^{n-1}(a_k - a_{k-1})(1+\epsilon)^k + a_0\right] \\ &= \varphi(1+\epsilon) > 0.\end{aligned}$$

This shows that on the circle $|z| = 1+\epsilon$, we have $|g(z)| < |f(z)|$, hence by Rouché's Theorem, we conclude that $f(z)$ and $f(z)+g(z) = (z-1)p(z)$ have the same number of n zeros inside the circle $|z| = 1+\epsilon$. Letting $\epsilon \to 0$ shows that all the zeros of $p(z)$ with $|z| \leq 1$ lie inside or on $|z| = 1$. Note that if $|z| = 1$ and $p(z) = 0$, then $f(z) = -g(z)$ and therefore $|g(z)| = |f(z)|$. But as we mentioned in the beginning of the proof, this is only possible when $z = 1$. Since $p(1) < 0$, all the $n-1$ roots of $p(z)$ with $|z| \leq 1$ lie inside $|z| < 1$. Together with the root $r > 1$ of $p(z)$, this accounts for all the n roots. □

Corollary 1. If $a_{n-1} \geq \cdots \geq a_1 \geq 1$ with $n > 1$, then the equation

$$p(x) := x^n - (a_{n-1}x^{n-1} + \cdots + a_1 x + 1) = 0$$

has a real root > 1 and $n-1$ roots of absolute value < 1.

Proof. This is clear. □

Theorem 3. Theorem 1 and Theorem 2 are equivalent.

Proof. In the first proof of Theorem 2, we show that Theorem 1 implies Theorem 2.

Now let's show the converse, i.e. Theorem 2 implies Theorem 1. Let $q(z) = a_n z^n + a_{n-1}z^{n-1} + \cdots + a_0$ be the given polynomial in Theorem 1, where $a_n > a_{n-1} > \cdots > a_0 > 0$. Consider an auxiliary polynomial

$$g(z) := (z-r)q(z) = a_n z^{n+1} + (a_{n-1} - ra_n)z^n + \cdots + (a_0 - ra_1)z - ra_0,$$

where r is a real number to be determined. Clearly the roots of polynomial $g(z)$ are the same as those of

$$\frac{g(z)}{a_n} = z^{n+1} - \left[\frac{ra_n - a_{n-1}}{a_n} \cdot z^n + \cdots + \frac{ra_1 - a_0}{a_n} z + \frac{ra_0}{a_n}\right],$$

where the coefficients of the polynomial in the brackets are positive if $r \geq 1$, and for $r = 1$, these coefficients sum up to 1. Now to apply Theorem 2, it suffices to choose $r > 1$ such that

$$ra_n - a_{n-1} \geq ra_{n-1} - a_{n-2} \geq \cdots \geq ra_1 - a_0 \geq ra_0$$

$$\Leftrightarrow r \geq \frac{a_k - a_{k-1}}{a_{k+1} - a_k} \text{ for } 0 \leq k \leq n-1,$$

where $a_{-1} = 0$. It follows that we may just choose any real number $r > 1$ such that

$$r \geq \max_{0 \leq k \leq n-1} \frac{a_k - a_{k-1}}{a_{k+1} - a_k}.$$

Then Theorem 2 says that there is a unique root of $g(z)$ greater than 1 (necessarily r) and all other roots are of absolute values less than 1. These other n roots are the zeros of $q(z)$, as required. □

3. Discussion

3.1. Applications of Theorem 2

The versatility of Theorem 2 allows it to bridge the gap between abstract mathematics and tangible real-world systems. Its applications are broad, spanning several critical fields:

3.1.1. Real-World Applications

- Population Dynamics: Utilizing the Leslie matrix to model age-structured growth and changes in biological populations.
- Finance: Calculating the Internal Rate of Return (IRR) to assess the profitability of potential investments.
- Reliability Engineering: Determining the Mean Time to Failure (MTTF) in survival analysis to predict system longevity.
- Information Technology: Powering search engine ranking algorithms (like PageRank) by analyzing the link structure of the web.
- Epidemiology: Modeling the spread of infectious diseases and predicting outbreak trajectories within populations.
- Control Theory: Ensuring system stability and optimizing feedback loops in engineering applications.

3.1.2. Theoretical Applications

Beyond practical implementation, Theorem 2 provides foundational insights into:

- Linear Recurrent Sequences: Analyzing the long-term behavior of sequences defined by linear recurrence relations.
- Number Theory: Exploring properties of algebraic integers and distribution patterns within numerical sets.

3.2. Future Research Directions

While Theorem 1 is supported by extensive literature, Theorem 2 – despite its prevalence in various contexts – has not yet been subjected to the same depth of analysis or generalization. This note establishes a formal bridge between the two theorems via suitable transformations. By establishing this link, we provide a framework to translate the well-documented generalizations of Theorem 1 into new, rigorous extensions for Theorem 2.

References

1. Eneström, G. Härledning af en allmän formel för antalet framtida dödsfall inom en afslutad grupp, med hänsyn tagen till den sannolika varaktigheten af de enskilda personernas lif. *Öfversigt af Kongliga Vetenskaps-Akademiens Förhandlingar.* 1893, 50, 405-415.

2. Kakeya, S. On the limits of the roots of an algebraic equation with positive coefficients. *Tôhoku Math. J. First Ser.* 1912, 2, 140-142.

3. Marden, M. Geometry of Polynomials (2nd ed. Mathematical Surveys No. 3). 1966, American Mathematical Society.

4. Hurwitz, A. Über einen Satz des Herrn Kakeya. *Tôhoku Math. J. First Ser.* 1913, 4, 89-93.

5. Anderson, N., Saff, E.B., and Varga, R.S. On the Eneström-Kakeya Theorem and its Sharpness. *Linear Algebra and its Applications.* 1979, 28, 5-16.

6. Eneström, G. Remarque sur un théorème relatif aux racines de l'équation $a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0$ où tous les coefficients sont réels et positifs. *Tôhoku Math. J.* 1920, 18, 34-36.