Fermat’s Spiral and the Line Between Yin and Yang

Taras Banakh, Oleg Verbitsky, and Yaroslav Vorobets

Abstract. Let \(D\) denote a disk of unit area. We call a set \(A \subset D\) perfect if it has measure \(1/2\) and, with respect to any reflection symmetry of \(D\), the maximal symmetric subset of \(A\) has measure \(1/4\). We call a curve \(\beta\) in \(D\) a yin-yang line if

\begin{itemize}
  \item \(\beta\) splits \(D\) into two congruent perfect sets,
  \item \(\beta\) crosses each concentric circle of \(D\) twice,
  \item \(\beta\) crosses each radius of \(D\) once.
\end{itemize}

We prove that Fermat’s spiral is the unique yin-yang line in the class of smooth curves algebraic in polar coordinates.

1. INTRODUCTION. The yin-yang concept comes from ancient Chinese philosophy. Yin and Yang refer to the two fundamental forces ruling everything in nature and human life. The two categories are opposite, complementary, and intertwined. They are depicted, respectively, as the dark and the light areas of the well-known yin-yang symbol (also Tai Chi or Taijitu; see Figure 1). The borderline between these areas represents in Eastern thought the equilibrium between Yin and Yang.

\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{yin_yang_symbols.png}
\caption{A gallery of yin-yang symbols: (a) a classical version, cf. [20]; (b) a modern version; (c) the version discussed in [12]; (d) an element of a Korean flag from the 19th century [17].}
\end{figure}

From the mathematical point of view, the yin-yang symbol is a bipartition of the disk by a certain curve \(\beta\), where by curve we mean the image of a real segment under an injective continuous map. We aim at identifying this curve and deriving an explicit mathematical expression for it. Such a project should apparently begin with choosing a set of axioms for basic properties of the yin-yang symbol in terms of \(\beta\). We begin with the following three.

\begin{itemize}
  \item \((A_1)\) \(\beta\) splits the disk \(D\) into two congruent parts.
  \item \((A_2)\) \(\beta\) crosses each concentric circle of \(D\) twice.
  \item \((A_3)\) \(\beta\) crosses each radius of \(D\) once (besides the center of \(D\), which must be visited by \(\beta\) due to \((A_2)\)).
\end{itemize}

While the first two properties are indisputable, our choice of the third axiom requires some motivation, as this condition is often not met in the actual yin-yang design.

doi:10.4169/000298910X521652
For example, in the most familiar versions some radii do not cross $\beta$ at all (Figures 1a and 1b). On the other hand, there are yin-yang patterns where the number of crossings varies from 1 to 2 (Figure 1d). Figure 1c shows an instance with $(A_3)$ obeyed, which arose from an analysis of the origins of the yin-yang symbolism [12]; see also [13, pp. 189–191].\textsuperscript{1} All the variations can be treated uniformly if we take a dynamic look at the subject. It will be deduced from our set of axioms that $\beta$ is a spiral. We therefore can consider a continuous family of yin-yang symbols determined by the spiral as shown in Figure 2. Thus, $(A_3)$ specifies a single representative that will allow us to expose the whole family.

![Figure 2](image)

**Figure 2.** Evolution of the yin-yang symbol (cf. Figure 1). Phase (c) is characterized by axioms $(A_1)$–$(A_6)$.

There is a huge variety of curves satisfying $(A_1)$–$(A_3)$ and hence a further specification is needed. We suggest another axiom, which is both mathematically beautiful and philosophically meaningful. It comes from the consideration of symmetry within $D$.

Denote the symmetry group of the disk $D$ by $\text{Sym}(D)$. As is well known, it consists of reflection and rotation symmetries. Focusing on these intrinsic symmetries of the disk, we will call a set $X \subseteq D$ symmetric if $s(X) = X$ for some nonidentity $s \in \text{Sym}(D)$.

Suppose now that $D$ has unit area. In fact, instead of area we will more often refer to the more general concept of Lebesgue measure. We call a set $A \subseteq D$ perfect if it has measure $1/2$ and any symmetric subset of $A$ has measure at most $1/4$. The latter is a kind of an extremal condition in view of the following fact: if an axis of symmetry $\ell$ of $D$ is drawn at random, then the largest subset of $A$ symmetric with respect to $\ell$ has measure $1/4$ on the average. From here it is not hard to derive a remarkable property of any perfect set: for every reflection symmetry $s \in \text{Sym}(D)$, the largest subset of

\textsuperscript{1}The border between Yin and Yang consists of two semicircles in Figure 1b and is Archimedes’ spiral in Figure 1c. In Figures 1a and 1d this is Fermat’s spiral, the subject of our further discussion.
A symmetric with respect to $s$ has measure exactly $\frac{1}{4}$. We prove these claims in Section 2. In the same section we demonstrate a simple example of a perfect set. The existence of such sets in the disk is really a magic fact: in $[3,4]$ we prove that perfect sets do not exist in the circle, nor in the sphere or the ball, even if higher dimensions are considered.

In words admitting profound esoteric interpretations, perfect sets demonstrate a sharp equilibrium between their “symmetric” and “asymmetric” parts, whatever particular reflection symmetry $s$ is considered. We put this phenomenon in our axiom system.

(A$_4$) $\beta$ splits $D$ into perfect sets (from now on it is supposed that $D$ has unit area).

There are a lot of examples of curves satisfying (A$_1$)–(A$_4$). Nevertheless, we are able to show that these four axioms determine a unique curve $\beta$ after imposing two other natural conditions.

(A$_5$) $\beta$ is smooth, i.e., has an infinitely differentiable parameterization $\beta : [0, 1] \to D$ with nonvanishing derivative.

The other condition expresses a belief that the border between Yin and Yang should be cognizable, i.e., it should be transcendent neither in the philosophical nor in the mathematical sense.

(A$_6$) $\beta$ is algebraic in polar coordinates.

By polar coordinates we mean the mapping $\Pi$ from the two-dimensional space of parameters $(\phi, r)$ onto the two-dimensional space of parameters $(x, y)$ defined by the familiar relations $x = r \cos \phi$ and $y = r \sin \phi$. The $(x, y)$-space is considered the standard Cartesian parameterization of the plane. Note that, like $x$ and $y$, both $\phi$ and $r$ are allowed to take on any real value. A curve in the $(\phi, r)$-plane is algebraic if all its points satisfy the equation $P(\phi, r) = 0$ for some bivariate nonzero polynomial $P$. A curve in the $(x, y)$-plane is algebraic in polar coordinates if it is the image of an algebraic curve under the mapping $\Pi$.

A classical instance of a curve both smooth and algebraic in polar coordinates is Fermat’s spiral (exactly this spiral is drawn in Figure 2). Fermat’s spiral is defined by the equation $a^2 r^2 = \phi$. The part of it specified by the restriction $0 \leq \phi \leq \pi$ (or, equivalently, $-\sqrt{\pi/a} \leq r \leq \sqrt{\pi/a}$) is inscribed in the disk of area $(\pi/a)^2$.

**Theorem 1.1.** Fermat’s spiral $\pi^2 r^2 = \phi$ is, up to congruence, the unique curve satisfying the axiom system (A$_1$)–(A$_6$).

Thus, if we are willing to accept axioms (A$_1$)–(A$_6$), the yin-yang symbol must look as in Figure 2c. Note that the factor of $\pi^2$ in the curve equation in Theorem 1.1 comes from the condition that $D$ has area 1. As all Fermat’s spirals are homothetic, we can equally well draw the yin-yang symbol using, say, the spiral $\phi = r^2$. Varying the range of $r$, we obtain modifications as twisted as desired. A Java applet and MetaPost code for drawing these curves can be found at [http://dx.doi.org/10.4169/loci003404](http://dx.doi.org/10.4169/loci003404).

Figure 3 shows a striking resemblance between Fermat’s spiral and the yin-yang line on old Korean flags. It is noteworthy that the flag depictions in Figure 3 are seemingly the oldest two known to date; see [21]. We are not aware of any historical background explaining this phenomenon.

In the next section we give background material on perfect sets. Theorem 1.1 is proved in Section 3. In Section 4 we discuss modifications of the axiom system (A$_1$)–
(\mathcal{A}_6). We observe that, after removal of (\mathcal{A}_5) or (\mathcal{A}_6), the system is no longer categorical, that is, does not specify the line between Yin and Yang uniquely. Finally, we mention an extension of Theorem 1.1 for a multipartite variant of the yin-yang symbol.

2. DISCLOSING SYMMETRY WITHIN THE DISK. We begin with mathematical motivation for (\mathcal{A}_4), the cornerstone of our yin-yang axiomatic system. The following Ramsey-type statement lies in the background: if we split the disk \( D \) into two parts in any way, at least one of them will contain a large symmetric subset. We will prove it in a more precise and stronger form. Recall that, throughout the paper, \( D \) denotes the disk of unit area.

**Theorem 2.1.** Any set \( A \subset D \) of measure 1/2 contains a symmetric subset of measure at least 1/4.

The theorem will follow from Lemma 2.2 below, which gives us even more detailed information. First, we need to fix some notation.

Given \( \psi \) in the range from 0 to 2\( \pi \), let \( R_\psi (r, \phi) = (r, \psi - \phi) \). In polar coordinates \((r, \phi)\), this describes the reflection of the plane in the line \( \phi = \psi/2 \). Let \( A \) be an arbitrary set of points in the plane. Note that any subset of \( A \) symmetric with respect to \( R_\psi \) is included in \( A \cap R_\psi A \) and the latter is itself symmetric with respect to \( R_\psi \).

If \( A \) is measurable, this observation allows us to speak of the Maximum Measure of a Symmetric SubSet of \( A \) with respect to \( R_\psi \), to be abbreviated as \( MS^A(\psi) \). Thus, we define

\[
MS^A(\psi) = \mu(A \cap R_\psi A),
\]

where \( \mu \), here and throughout, stands for Lebesgue measure.
Lemma 2.2. Let $S$ be the circle of unit length and $D$ be the disk of unit area. Suppose that $S$ and $D$ are concentric and put the origin of polar coordinates at the center. We will consider measurable subsets of $S$ or $D$. In the former case $\mu$ will refer to Lebesgue measure on $S$; in the latter case $\mu$ will be Lebesgue measure in the plane. Let $\psi$ be a random angle.

1. If $A \subseteq S$, then the average value of $M^A_S(\psi)$ is equal to $\mu(A)^2$.

2. If $A \subseteq D$, then the average value of $M^A_S(\psi)$ is greater than or equal to $\mu(A)^2$.

Proof. 1. Given a set $U \subseteq S$, let $\chi_U$ denote the characteristic function of $U$. More precisely, $\chi_U(\phi) = 1$ if the point of $S$ with angular coordinate $\phi$ is in $U$ and $\chi_U(\phi) = 0$ otherwise. Note that

$$M^A_S(\psi) = \frac{1}{2\pi} \int_0^{2\pi} \chi_{A \cap R_\psi A}(\phi) \, d\phi \quad \text{and} \quad \chi_{A \cap R_\psi A} = \chi_A(\phi) \chi_A(\psi - \phi).$$

Averaging $M^A_S(\psi)$ on $\psi$ and changing the order of integration, we obtain

$$\frac{1}{2\pi} \int_0^{2\pi} M^A_S(\psi) \, d\psi = \frac{1}{(2\pi)^2} \int_0^{2\pi} \chi_A(\phi) \int_0^{2\pi} \chi_A(\psi - \phi) \, d\psi \, d\phi = \frac{1}{(2\pi)^2} \left( \int_0^{2\pi} \chi_A(\phi) \, d\phi \right)^2 = \mu(A)^2.$$

2. Let $S_r$ denote the concentric circle of $D$ having radius $r$ and $A_r = A \cap S_r$. On $S_r$ we consider the measure $\mu_r$ normed so that $\mu_r(S_r) = 1$. Note that

$$M^A_S(\psi) = \int_0^{1/\sqrt{\pi}} 2\pi r \mu_r(A_r \cap R_\psi A_r) \, dr,$$

where $A_r$ is measurable for almost all $r$ by Fubini’s theorem. Averaging $M^A_S(\psi)$ on $\psi$, changing the order of integration, and using part 1, we have

$$\frac{1}{2\pi} \int_0^{2\pi} M^A_S(\psi) \, d\psi = \int_0^{1/\sqrt{\pi}} \frac{1}{2\pi} \int_0^{2\pi} \mu_r(A_r \cap R_\psi A_r) \, d\psi \, 2\pi r \, dr = \int_0^{1/\sqrt{\pi}} 2\pi r \mu_r(A_r)^2 \, dr.$$

The proof is finished by an application of the Cauchy-Schwarz inequality. Recall that this inequality says that

$$\int a(r)^2 \, dr \int b(r)^2 \, dr \geq \left( \int a(r)b(r) \, dr \right)^2.$$

Using the equality

$$\int_0^{1/\sqrt{\pi}} 2\pi r \, dr = 1,$$
\[
\frac{1}{2\pi} \int_0^{2\pi} M S^A(\psi) \, d\psi = \int_0^{1/\sqrt{\pi}} \left( \sqrt{2\pi r \mu_r(A_r)} \right)^2 \, dr \int_0^{1/\sqrt{\pi}} \left( \sqrt{2\pi r} \right)^2 \, dr \\
\geq \left( \int_0^{1/\sqrt{\pi}} 2\pi r \mu_r(A_r) \, dr \right)^2 = \mu(A)^2.
\]

Part 2 of Lemma 2.2 immediately implies that, for any measurable set \( A \subseteq D \), we must have \( M S^A(\psi) \geq \mu(A)^2 \) for at least one \( \psi \). If \( \mu(A) = 1/2 \), this is exactly the statement of Theorem 2.1. We will be interested in cases showing that the estimate of Theorem 2.1 is sharp.

**Definition 2.3.** Let \( D \) denote the disk of unit area. Call a set \( A \subset D \) perfect if \( \mu(A) = 1/2 \) and every symmetric subset of \( A \) has measure at most \( 1/4 \).

**Lemma 2.4.** If a set \( A \subset D \) is perfect, then \( M S^A(\psi) = 1/4 \) for all \( \psi \).

**Proof.** By the definition of a perfect set, \( M S^A(\psi) \leq 1/4 \) for every \( \psi \). Combining this with part 2 of Lemma 2.2, we conclude that the average value of \( M S^A(\psi) \) is exactly equal to \( 1/4 \) and, furthermore, \( M S^A(\psi) = 1/4 \) for almost all \( \psi \). Equality actually holds true for all \( \psi \) because \( M S^A(\psi) \) is a continuous function of \( \psi \). We leave the latter claim as an exercise for the interested reader. (Hint: Prove it first for compact \( A \); then use the fact that measurable sets are approximable by compacta.)

At first glance, the property of a perfect set established in Lemma 2.4 may appear rather restrictive, and a priori it is not so clear whether this notion is consistent. Let us remove all doubts.

**Theorem 2.5.** Perfect sets exist.

We precede the proof of Theorem 2.5 with a technical suggestion that will also be useful while proving our main result in the next section. Our goal is to transform the disk into a cylinder so that measure and symmetry are preserved. Let \( \mathbb{S} \) be the circle of unit length. Consider the cylinder \( C = \mathbb{S} \times (0, 1] \) supplied with the product measure \( \mu = \lambda \times \lambda \), where \( \lambda \) is Lebesgue measure. A point in \( C \) will be described by a pair of coordinates \((u, v)\) with \( 0 < u \leq 1 \) and \( 0 < v \leq 1 \). In this way \( C \) is identified with its development onto the square \((0, 1] \times (0, 1] \). For the disk \( D \) we use polar coordinates \((r, \phi)\) with \( 0 < r \leq \pi^{-1/2} \) and \( 0 < \phi \leq 2\pi \). Delete the center \( O \) from \( D \) and define a one-to-one mapping \( T: D \setminus \{O\} \to C \) by

\[
T(r, \phi) = \left( \frac{\phi}{2\pi}, \pi r^2 \right).
\]

In \( \mathbb{S} = (0, 1] \) we will perform addition and subtraction as in the additive group \( \mathbb{R}/\mathbb{Z} \).

Since \( T \) is a diffeomorphism, a set \( X \subseteq D \) is measurable if and only if \( T(X) \) is. To see the geometric meaning of \( T \), notice that it takes a radius of the disk onto a longitudinal section of the cylinder. Furthermore, a concentric circle of \( D \) is taken onto a cross section of \( C \) so that the area below the section is equal to the area within the circle. It follows that, if a set \( X \subseteq D \) is measurable, then \( X \) and \( T(X) \) have the same measure.
The transformation $T$ preserves symmetry in the following sense: for every $s \in \text{Sym}(D)$ there is an $s' \in \text{Sym}(C)$ such that, for any $X \subseteq D$, we have $s(X) = X$ if and only if $s'(T(X)) = T(X)$. Specifically, the rotation $s'(u, v) = (u + h, v)$ of the cylinder corresponds to the rotation of the disk by the angle $2\pi h$, and the reflection $s'(u, v) = (g - u, v)$ of the cylinder corresponds to the reflection of the disk in the line $\phi = \pi g$.

Proof of Theorem 2.5. Making use of the transformation $T$, we prefer to deal with the cylinder $C$. Let $A \subset C$ be as shown in Figure 4 on the left side. We claim that $T^{-1}(A)$ is perfect. It suffices to verify the following three properties of $A$:

- $\mu(A) = 1/2$ (which is evident).
- $A$ contains no subset symmetric with respect to any rotation. Indeed, whatever point $a \in A$ and rotation $s$ of $C$ are considered, the orbit of $a$ with respect to $s$ cannot stay completely inside $A$.
- Every subset of $A$ symmetric with respect to a reflection has measure at most $1/4$. Let $B = (0, \frac{1}{2}]$, which is the projection of one of the two components of $A$ onto the circle $S$. Consider a reflection $s_g(u, v) = (g - u, v)$. For the largest subset of $A$ symmetric with respect to $s_g$ we have

$$
\mu(A \cap s_g A) = \frac{1}{2} \lambda(B \cap (g - B)) + \frac{1}{2} \lambda\left(\left(\frac{1}{4} + B\right) \cap \left(g - \frac{1}{4} - B\right)\right).
$$

Note that the latter measure is equal to

$$
\lambda\left(\left(\frac{1}{2} + B\right) \cap (g - B)\right) = \lambda\left((S \setminus B) \cap (g - B)\right)
$$

and, therefore,

$$
\mu(A \cap s_g A) = \frac{1}{2} \lambda(g - B) = \frac{1}{4}.
$$

Figure 4. Construction of a perfect set (the disk on the right side and the cylinder development on the left side).

The construction presented in the proof of Theorem 2.5 can easily be modified to give a variety of other examples of perfect sets. Essentially the same argument goes through whenever a set $A$ has two properties. First, every cross section of $A$ has to be a semicircle. Second, if we split $A$ into two components by the cross section $v = \frac{1}{2}$, then the upper component has to be obtainable from the lower one by lifting it and rotating by the angle $\pi/2$. The latter condition plays an important role in the proof of our main result and will be referred to as the translation property.
3. ESTABLISHING THE BORDERLINE BETWEEN YIN AND YANG. We begin with a brief overview of the proof of Theorem 1.1. Using the mapping $T$ from Section 2, we will transform the disk (minus the center) into a cylinder, which then will be unwrapped into the plane. In this way a curve $\beta$ (more precisely, one of its two branches emanating from the center) will be transformed into a curve $\alpha$ in the unit square. In particular, the Fermat spiral is converted in this way into a straight line segment. Our goal becomes to show that the line segment is the only curve in the unit square satisfying the properties implied by axioms $(A_1)$–$(A_6)$.

The proof of Theorem 1.1 consists essentially of two parts. The first is related to our construction of a perfect set in Theorem 2.5. At the end of Section 2 we formulated the translation property, which is a sufficient condition for perfectness. Now we will prove that this property is also a necessary condition. In terms of $\alpha$, this means that the line $v = \frac{1}{2}$ splits $\alpha$ into two congruent parts (see Lemma 3.1). As a consequence, some translation of $\alpha$ intersects $\alpha$ infinitely many times. This condition is obviously met if $\alpha$ is a line segment. In the second part of the proof we show that any other possibility is excluded by axioms $(A_5)$ and $(A_6)$ (see Lemma 3.2), that is, $\alpha$ must be a line segment and $\beta$ must be Fermat’s spiral.

3.1. Proof of Theorem 1.1. Assume that $\beta$ satisfies axioms $(A_1)$–$(A_3)$. As was already said, we take into consideration another curve $\alpha$ associated with $\beta$ in a natural way. It easily follows from $(A_1)$ and $(A_2)$ that $\beta$ is symmetric with respect to the center $O$ of the disk $D$ and that $O$ splits $\beta$ into two congruent branches, say, $\beta_1$ and $\beta_2$. We will use the transformation $T : D \setminus \{O\} \to (0, 1]^2$ described before the proof of Theorem 2.5, where the square $S \times (0, 1]$ is thought of as the development of the cylinder $C = S \times (0, 1]$. Let $\alpha$ be the image of $\beta_1$ under $T$. By $(A_3)$, this curve can be considered to be the graph of a function $v = \alpha(u)$, where $u$ ranges over a half-open interval of length $1/2$. Shifting, if necessary, the coordinate system and appropriately choosing the positive direction of the $u$-axis, we suppose that $\alpha$ is defined for $0 < u \leq \frac{1}{4}$ and $\lim_{u \to 0} \alpha(u) = 0$. Since $\beta_1$ goes from the center to a peripheral point, we have $\alpha(1/2) = 1$. Since $\beta_1$ crosses each concentric circle exactly once, $\alpha$ is a monotonically increasing function. Note that the image of $\beta_2$ under $F$ is obtained from $\alpha$ by translation by $1/2$, that is, it is the graph of the function $v = \alpha(u - 1/2)$ on $1/2 < u \leq 1$ (see the left square in Figure 5).

![Figure 5](image.png)

**Figure 5.** Proof of Lemma 3.1: regions $A$ and $\bar{A}$ in the cylinder $C$.

**Lemma 3.1.** Suppose that $\beta$ satisfies $(A_1)$–$(A_3)$. Let $\alpha$ be the associated curve. Then $\beta$ satisfies $(A_4)$ if and only if

$$\alpha \left( u + \frac{1}{4} \right) = \alpha(u) + \frac{1}{2} \quad \text{whenever} \quad 0 < u \leq \frac{1}{4}. \quad (1)$$
Proof. The curve $\alpha$ and its shift by the angle $\pi$ split the cylinder $C$ into two congruent parts, say, $A$ and $B$. They are the images under $T$ of the two parts into which $D$ is split by $\beta$.

If equality (1) holds true, then $A$ has the translation property discussed at the end of Section 2, and the perfectness of $T^{-1}(A)$ and $T^{-1}(B)$ can be proved by the method used to prove Theorem 2.5. We leave details to the reader.

Conversely, assume $(A_4)$. Consider the development of the cylinder onto the square and suppose that $A$ is the region in $(0, 1]^2$ between $\alpha$ and its shift by $1/2$. We first introduce some notation. Given $v \in (0, 1]$, let $A_v = \{u : (u, v) \in A\}$. Note that $A_v$ is the interval with endpoints $\alpha^{-1}(v)$ and $\alpha^{-1}(v) + 1/2$; we will sometimes write $A_v = \alpha^{-1}(v) + H$, where $H = (0, 1/2)$.

Furthermore, let $\bar{\alpha}(u) = \alpha(u/2)$, for $0 < u \leq 1$, and let $\bar{A}$ be the region of the cylinder $C$ bounded by the graph of $\bar{\alpha}$ and its image under rotation by the angle $\pi$. Note that $\bar{A}$ is obtained from $A$ by torsion of the cylinder $C$, and in the development of $C$ this region looks like a stretching of $A$; see the right square in Figure 5. Given $u \in (0, 1]$, let $\bar{A}_u = \{v : (u, v) \in \bar{A}\}$ and $m(u) = \lambda(\bar{A}_u)$. If $0 < u \leq 1/2$, we have

$$m(u) = \bar{\alpha}(u) + \left(1 - \bar{\alpha}\left(u + \frac{1}{2}\right)\right), \quad (2)$$

$$m \left(u + \frac{1}{2}\right) = \bar{\alpha} \left(u + \frac{1}{2}\right) - \bar{\alpha}(u).$$

We are now prepared to prove (1). Since $A$ is the image of a perfect set, by Lemma 2.4 we have

$$\mu(A \cap s_g A) = \frac{1}{4} \quad (3)$$

for any reflection $s_g(u, v) = (g - u, v)$. Here $\mu = \lambda \times \lambda$ is the measure on $(0, 1]^2$. Now we come to the main point of the proof:

$$\mu(A \cap s_g A) = \int_{g-1/2}^{g} m(u) \, du \quad \text{whenever } \frac{1}{2} < g \leq 1. \quad (4)$$

Assume (4) for the moment. It follows by (3) that the expression $\int_{g-1/2}^{g} m(u) \, du$ does not depend on $g$. Differentiating it as a function of $g$, we arrive at the identity $m(g) - m(g - 1/2) = 0$, that is,

$$m(u) = m \left(u + \frac{1}{2}\right) \quad \text{whenever } 0 < u \leq \frac{1}{2}. \quad (5)$$

Substituting from (2), we get

$$2\bar{\alpha} \left(u + \frac{1}{2}\right) = 2\bar{\alpha}(u) + 1.$$
In particular, for $0 < u \leq 1/4$ we have
\[ \tilde{\alpha} \left( 2u + \frac{1}{2} \right) = \tilde{\alpha}(2u) + \frac{1}{2}, \]
which in terms of $\alpha$ is exactly what is claimed in (1).

To complete the proof, it remains to prove (4). Observe that
\[ \lambda(A_v \cap s_g A_v) = \lambda((\alpha^{-1}(v) + H) \cap (-\alpha^{-1}(v) + g - H)) \]
\[ = \lambda((2\alpha^{-1}(v) + H) \cap (g - H)) = \lambda((\tilde{\alpha}^{-1}(v) + H) \cap (g - H)) \]
for any $g$. If $1/2 < g \leq 1$, we can continue this chain of equalities as follows:
\[ \lambda(A_v \cap s_g A_v) = \int_{g-1/2}^{g} \chi_{\tilde{\alpha}^{-1}(v)+H}(u) \, du \]
\[ = \int_{g-1/2}^{g} \chi_{\tilde{\alpha}}(u, v) \, du = \int_{g-1/2}^{g} \chi_{\tilde{\alpha}u}(v) \, du. \]

Finally,
\[ \mu(A \cap s_g A) = \int_0^1 \lambda(A_v \cap s_g A_v) \, dv = \int_0^1 \int_{g-1/2}^{g} \chi_{\tilde{\alpha}u}(v) \, du \, dv \]
\[ = \int_{g-1/2}^{g} \int_0^1 \chi_{\tilde{\alpha}u}(v) \, dv \, du = \int_{g-1/2}^{g} m(u) \, du, \]
as needed. \[\blacksquare\]

The simplest curve satisfying (1) is the line $v = 2u$, $0 \leq u \leq 1/2$. As is easily seen, under the application of $T^{-1}$ this line and its shift $v = 2u - 1$, $1/2 \leq u \leq 1$, are transformed into Fermat’s spiral $\pi^2 r^2 = \phi$, $0 \leq \phi \leq \pi$. It follows that Fermat’s spiral, which clearly satisfies $(A_1)$–$(A_3)$, also satisfies $(A_4)$. Note that $(A_5)$ and $(A_6)$ hold true as well: Fermat’s spiral is explicitly given by an algebraic relation in polar coordinates and is smooth, for example, due to the parameterization
\[
\begin{align*}
  x &= \frac{1}{\pi} t \cos t^2, \\
  y &= \frac{1}{\pi} t \sin t^2.
\end{align*}
\]

To prove Theorem 1.1, we hence have to prove that, if a curve $\beta$ satisfies $(A_1)$–$(A_6)$, then it coincides, up to congruence, with Fermat’s spiral. It suffices to prove that the associated curve $\alpha$ is a straight line segment.

Since $\beta$ is smooth by $(A_3)$ and $\alpha$ is a smooth transformation of a part of $\beta$, we conclude that $\alpha$ is smooth as well. As follows from $(A_6)$ and the definition of the transformation $T$, all points of $\alpha$ satisfy the relation $P(u, \sqrt{v}) = 0$ for a nonzero bivariate polynomial $P$. Since this relation can be rewritten in the form $\sqrt{v} Q(u, v) + R(u, v) = 0$ for some polynomials $Q$ and $R$, the points of $\alpha$ satisfy the polynomial relation $\sqrt{v} Q^2(u, v) - R^2(u, v) = 0$.

Thus, $\alpha$ is a smooth algebraic curve. By Lemma 3.1, $\alpha$ satisfies relation (1). This means that $\alpha$ has infinite intersection with its shift $\alpha + (\frac{1}{4}, \frac{1}{2})$. Theorem 1.1 immediately follows from the following fact.

**Lemma 3.2.** If a smooth algebraic curve $\gamma \subset \mathbb{R}^2$ has infinite intersection with its shift by a nonzero vector $(a, b)$, then $\gamma$ is contained in a line.

November 2010] FERMAT’S SPIRAL 795
The proof of Lemma 3.2 takes the rest of this section.

3.2. Background on Smooth Algebraic Curves. In what follows we suppose that a curve \( \gamma \subseteq \mathbb{R}^2 \) is the image of a real open interval under an injective continuous map (the forthcoming considerations apply to \( \alpha \), which is the image of a half-open interval, after removal from \( \alpha \) of its endpoint). Recall that \( \gamma \) is smooth if it has an infinitely differentiable parameterization \( \gamma : (0, 1) \rightarrow \mathbb{R}^2 \) with nonvanishing derivative. A curve \( \gamma \) is algebraic if every point \( (x, y) \in \gamma \) satisfies the relation \( P(x, y) = 0 \) for a nonzero polynomial \( P \) with real coefficients. The degree of \( \gamma \) is the minimum possible degree of such a \( P \). If \( P \) is irreducible, i.e., does not admit a factorization \( P(x, y) = Q(x, y)R(x, y) \) with \( Q \) and \( R \) both being nonconstant real polynomials, then \( \gamma \) is called irreducibly algebraic.

Theorem 3.3 (Bézout, see, e.g., [11]). Suppose that \( P_1 \) and \( P_2 \) are distinct bivariate polynomials of degree \( m \) and \( n \), respectively. Let \( \rho_i = \{(x, y) \in \mathbb{R}^2 : P_i(x, y) = 0\} \), \( i = 1, 2 \). If both \( P_1 \) and \( P_2 \) are irreducible, then \( \rho_1 \) and \( \rho_2 \) have at most \( mn \) points in common.

A function \( f : \mathbb{R} \rightarrow \mathbb{R} \) is called analytic at a point \( x_0 \) if in some neighborhood of \( x_0 \) the function can be developed in a power series

\[
f(x) = \sum_{n=0}^{\infty} a_n(x - x_0)^n
\]

with real coefficients \( a_n \). Furthermore, we call \( f \) analytic on an open set \( U \) if it is analytic at every \( x_0 \in U \). It can be showed that, if \( f \) is analytic at a point \( x_0 \), then it is analytic on a neighborhood of \( x_0 \).

Theorem 3.4 (The uniqueness theorem, see, e.g., [15]). Let functions \( f \) and \( g \) be analytic on a connected set \( U \). Suppose that the set \( \{x \in U : f(x) = g(x)\} \) has an accumulation point inside \( U \). Then \( f(x) = g(x) \) everywhere on \( U \).

Finally, we introduce a useful concept stronger than that of a smooth curve. We call a curve \( \gamma \subseteq \mathbb{R}^2 \) analytic if for each point \( (x_0, y_0) \in \gamma \) there is a neighborhood of \( (x_0, y_0) \) such that within this neighborhood either \( \gamma \) is the graph of a real function \( y = f(x) \) analytic at \( x_0 \), or it is the graph of a real function \( x = f(y) \) analytic at \( y_0 \).

3.3. Proof of Lemma 3.2. We begin with a simpler particular case. Assume that \( \gamma \) is irreducibly algebraic.

Let \( P(x, y) \) be an irreducible polynomial such that \( \gamma \subseteq \rho \), where \( \rho = \{(x, y) \in \mathbb{R}^2 : P(x, y) = 0\} \). Note that \( \rho + (a, b) = \{(x, y) \in \mathbb{R}^2 : P(x - a, y - b) = 0\} \) and that the polynomial \( P(x - a, y - b) \) is irreducible as well. Since the intersection \( \rho \cap (\rho + (a, b)) \supseteq \gamma \cap (\gamma + (a, b)) \) is infinite, Bézout’s theorem implies the equality \( \rho = \rho + (a, b) \). It follows that \( \rho \) has infinite intersection with the line \( \ell = \{(x, y) \in \mathbb{R}^2 : b(x - x_0) - a(y - y_0) = 0\} \) for an arbitrary \( (x_0, y_0) \in \rho \). Since any linear polynomial is irreducible, we can apply Bézout’s theorem once again and conclude that \( \ell = \rho \supseteq \gamma \).

To prove Lemma 3.2, it now suffices to show that every smooth algebraic curve is irreducibly algebraic. This can be done in two steps, corresponding to our next two lemmas.
Lemma 3.5. Every smooth algebraic curve is analytic.

The implicit function theorem implies that an algebraic curve is analytic everywhere except a finite number of so-called singular points. Thus, Lemma 3.5 essentially asserts that smoothness rules out singularity. Though this sounds like a tautology, an accurate proof requires some tools from the theory of analytic functions. We omit further details, noting only that Lemma 3.5 can be deduced, e.g., from [9, Theorem 6].

The next lemma completes the reduction to the irreducible case and finishes the proof of Theorem 1.1.

Lemma 3.6. Every analytic algebraic curve \( \gamma \subset \mathbb{R}^2 \) is irreducibly algebraic.

Proof. Let \( P(x, y) \) be a nonzero polynomial of the smallest possible degree such that \( \gamma \subseteq \{(x, y) \in \mathbb{R}^2 : P(x, y) = 0\} \). We will prove that \( P(x, y) \) is irreducible.

Assume to the contrary that \( P(x, y) = P_1(x, y)P_2(x, y) \) with \( P_1 \) and \( P_2 \) being nonconstant polynomials. For \( i = 1, 2 \) let \( \gamma_i = \{(x, y) \in \gamma : P_i(x, y) = 0\} \) and \( \gamma_i' \) be the set of nonisolated points of \( \gamma_i \). It is clear that the sets \( \gamma_i \) and \( \gamma_i' \) are closed in \( \gamma \). We claim that \( \gamma_i' \) is also open in \( \gamma \).

Indeed, take an arbitrary point \( (x_0, y_0) \in \gamma_i' \). Using the analyticity of \( \gamma \) and swapping the roles of \( x \) and \( y \) if necessary, we can assume that in some neighborhood of \( (x_0, y_0) \) the curve \( \gamma \) coincides with the graph of a real function \( f(x) \) analytic on an interval \( (x_0 - \varepsilon, x_0 + \varepsilon) \). Notice that the function \( g_i(x) = P_i(x, f(x)) \) is analytic on the same interval. Denote the projection of \( \gamma_i \) onto the \( x \)-axis by \( \eta_i \). Since \( (x_0, y_0) \in \gamma_i' \), the set \( \eta_i \cap (x_0 - \varepsilon, x_0 + \varepsilon) \) is infinite and contains \( x_0 \) as an accumulation point. By the uniqueness theorem, \( g_i(x) = 0 \) for all \( x \in (x_0 - \varepsilon, x_0 + \varepsilon) \). Consequently, \( \{(x, f(x)) : |x - x_0| < \varepsilon\} \subset \gamma_i \), and hence \( (x_0, y_0) \) belongs to the interior of \( \gamma_i' \) in \( \gamma \).

Thus, both the sets \( \gamma_i' \) and \( \gamma_i' \) are closed and open in \( \gamma \). Since \( \gamma = \gamma_1' \cup \gamma_2' \), one of the sets is nonempty and hence coincides with \( \gamma \) by connectedness of \( \gamma \). We see that \( \gamma = \gamma_i \) is included in \( \{(x, y) \in \mathbb{R}^2 : P_i(x, y) = 0\} \) for some \( i \). This is a contradiction because \( P_i \) has degree strictly less than the degree of \( P \).

4. MODIFICATIONS OF THE AXIOM SYSTEM \((A_1)\)–\((A_6)\). We may wonder what happens when we deviate a little from our set of axioms. An important observation is that, if either \((A_5)\) or \((A_6)\) is removed, the yin-yang borderline is no longer specified uniquely by the relaxed axiom system. We also discuss multipartite variants of the yin-yang symbol.

4.1. Smoothness and Algebraicity Are Essential. The following example demonstrates that there are a variety of noncongruent curves \( \beta \) satisfying axioms \((A_1)\)–\((A_5)\). By Theorem 1.1, all such curves excepting Fermat’s spiral fail to meet \((A_6)\).

Example 4.1. Let \( 0 < \lambda < 1/4 \). The polar equation

\[
\pi^2 r^2 = \phi + \lambda \sin 4\phi, \quad \text{where} \ 0 \leq \phi \leq \pi \ (\text{or} \ -\pi^{-1/2} \leq r \leq \pi^{-1/2}),
\]

determines an analytic curve \( \beta \) satisfying \((A_1)\)–\((A_3)\).

Now we exhibit noncongruent curves satisfying \((A_1)\)–\((A_4)\) and \((A_6)\) that are differentiable any preassigned number of times.
Example 4.2. Let $\lambda$ be a positive real number and $k$ a nonnegative integer. Define a function

$$f(u) = \begin{cases} 
  f_1(u) = 2u + \lambda u^{k+1} \left( \frac{1}{4} - u \right)^{k+1}, & 0 \leq u \leq \frac{1}{4}, \\
  f_2(u) = \frac{1}{2} + f_1 \left( u - \frac{1}{4} \right), & \frac{1}{4} \leq u \leq \frac{1}{2}.
\end{cases}$$

For each sufficiently small $\lambda$, the polar equation

$$\pi r^2 = f \left( \frac{\phi}{2\pi} \right), \text{ where } 0 \leq \phi \leq \pi \text{ (or } -\pi^{-1/2} \leq r \leq \pi^{-1/2}),$$

determines a $k$-times continuously differentiable curve $\beta$ satisfying axioms $(A_1)$–$(A_4)$ and $(A_6)$.

Verification of Examples 4.1 and 4.2 is routine and left to the interested reader.

4.2. Multipartite Variants. For the purpose of the current discussion, we will call a curve a spiral if it is described by a polar equation

$$|r| = f(\phi), \quad 0 \leq \phi \leq \pi \ell,$$

where $f$ is a strictly increasing real function. We will say that the spiral makes $\ell$ turns.

3-part yin-yang symbols are frequently seen in Korea, where they are called Sam-Taegeuk, and in Zen temples in the Himalayas, where they are called Gankyil or Wheel of Joy (see Figure 6a). The Gankyil has a 4-part variant [5]. From the mathematical point of view, a $k$-part yin-yang symbol is a partition of the unit disk $D$ by a $k$-coil spiral into $k$ congruent parts. Equivalently, we can speak of a 1-coil spiral $\sigma$ along with its rotations by angles $2\pi i/k$ for all $i < k$. Here 1-coil means that only positive values of $r$ are allowed. Extending the notion of a perfect set to any measurable set $A \subset D$, we call $A$ perfect if every symmetric subset of $A$ has measure at most $\mu(A)^2$ (symmetric subsets of measure at least $\mu(A)^2$ always exist by Lemma 2.2). We adapt our axiomatic system in the $k$-partite case by requiring that each part of the symbol be perfect and that $\sigma$ be smooth and algebraic in polar coordinates. Our main result, Theorem 1.1, generalizes to the $k$-partite case as follows: among 1-turn spirals, only Fermat’s spiral meets the postulated conditions (see Figures 6b and 6c).
4.3. Related Work. The present paper falls naturally into the context of the branch of Ramsey theory studying the persistence of symmetry under “arbitrarily bad” partitions of various structures. Other results and open problems in this research area can be found in [1, 2, 3, 4, 14].

The semicircular yin-yang line became widespread as recently as in the 20th century; see [13, p. 189]. Bisection of this popular diagram is a known puzzle in recreational mathematics [8, 10, 16]. Graphical aspects of the yin-yang design, including 3-dimensional variants, are discussed in [6, 7].

ACKNOWLEDGMENTS. We thank Igor Chyzhykov for referring us to the book [9], Maria Korolyuk for useful comments, Iryna Vus for her kind assistance with graphics, and an anonymous referee for many valuable expository suggestions. Our special thanks to Mihyun Kang, Kyoko Slany, and Kaori Yamazaki for their helpful guidelines on the history of the Korean flag and Japanese sources on it. The second author was in part supported by the Alexander von Humboldt foundation.

REFERENCES

1. T. Banakh, B. Bokalo, I. Guran, T. Radul, and M. Zarichnyi, Problems from the Lviv topological seminar, in Open Problems in Topology II, E. Pearl, ed., Elsevier, Amsterdam, 2007, 655–667.
2. T. Banakh and I. Protasov, Symmetry and colorings: Some results and open problems, Izvestiya Gomel'skogo Universiteta. Voprosy Algebra 4(17) (2001) 5–16; also available at http://arxiv.org/abs/0901.3356.
3. T. Banakh, Ya. Vorobets, and O. Verbitsky, Ramsey-type problems for spaces with symmetry, Izvestiya RAN, seriya matematicheskaya 64(6) (2000) 3–40; English translation from Russian in Russian Academy of Sciences. Izvestiya Mathematics 64 (2000) 1091–1127. doi:10.1070/IM2000v064n06ABEH000310
4. ———, A Ramsey treatment of symmetry, Electron. J. Combin. 7 (2000) R52.
5. R. Beer, The Handbook of Tibetan Buddhist Symbols, Serindia Publications, Chicago, 2003.
6. C. Browne, Taiji variations: Yin and Yang in multiple dimensions, Computers & Graphics 31 (2007) 142–146. doi:10.1016/j.cag.2006.10.005
7. C. Browne and P. Wamelen, Spiral packing, Computers & Graphics 30 (2006) 834–842. doi:10.1016/j.cag.2006.07.010
8. H. E. Dudeney, Amusements in Mathematics, Dover, Mineola, NY, 1970.
9. B. A. Fuks and V. I. Levin, Functions of a Complex Variable and Their Applications, GITTL, Moscow, 1951; English translation from Russian by J. Berry, Functions of a Complex Variable and Some of Their Applications II, Pergamon Press, Oxford, 1961.
10. M. Gardner, My Best Mathematical and Logic Puzzles, Dover, Mineola, NY, 1994.
11. F. Kirwan, Complex Algebraic Curves, Cambridge University Press, Cambridge, 1992.
12. S. Li, On the relationship between the shaping of taijitu picture and ancient astronomical observation (in Chinese), Dongnan Wenhua (Culture of Southeast China) 1991.03/04.
13. F. Louis, The genesis of an icon: The taiji diagram’s early history, Harvard Journal of Asiatic Studies 63 (2003) 145–196. doi:10.2307/25066694
14. G. Martin and K. O’Bryant, The symmetric subset problem in continuous Ramsey theory, Experiment. Math. 16 (2007) 145–166.
15. R. Narasimhan and Y. Nievergelt, Complex Analysis in One Variable, 2nd ed., Birkhäuser, Boston, MA, 2001.
16. C. W. Trigg, Bisecting of Yin and of Yang, Math. Mag. 34 (1960) 107–108.
17. Wikipedia Commons, http://en.wikipedia.org/wiki/Image:Taegukgi.jpg.
18. ———, http://en.wikipedia.org/wiki/Image:Flag_of_old_Korea.jpg.
19. ———, http://en.wikipedia.org/wiki/File:Sam_Taeguk.jpg.
20. Yin-yang, Micropædia, vol. 12, The New Encyclopaedia Britannica, 15th ed., Encyclopaedia Britannica, Chicago, 1991, p. 845.
21. S. Yu, Discovery of old flag discredits “Taegukgi” legend, The Chosun Ilbo, January 26, 2004, available at http://english.chosun.com/w21data/html/news/200401/200401260030.html.

TARAS BANAKH is Professor of Mathematics at Universities of Lviv (Ukraine) and Kielce (Poland). He can serve as a confirmation of the birthday paradox, sharing a birthday on March 30 with Stefan Banach, the most famous mathematician who has ever worked at Lviv University. Moreover, they are namesakes in Ukrainian:
Тарас Банах and Стефан Банах, respectively. One may speculate about whether this is a reason why T. Banakh reflects not only on problems of topology and Banach space theory but also on the problem of reincarnation and the mystery of the yin-yang symbol.

Department of Mechanics and Mathematics, Lviv National University, Universytetska 1, Lviv 79000, Ukraine, and

Uniwersytet Humanistyczno-Przyrodniczy Jana Kochanowskiego w Kielcach, Poland
tbanakh@yahoo.com

OLEG VERBITSKY studied mathematics first at Lviv University and then at Moscow State University, where he earned his Ph.D. under the direction of Sergei Adian and Alexander Razborov in 1994. Since then he has alternated teaching at Lviv University while holding the position of senior researcher at the Institute of Applied Problems of Mechanics and Mathematics in Lviv with leading the life of a traveling researcher (in particular, as a Lise Meitner fellow at Vienna University of Technology and a Humboldt fellow at Humboldt University of Berlin). He works in complexity theory, also maintaining research interests in combinatorics and logic.

Institute for Applied Problems of Mechanics and Mathematics, Naukova St. 3-Б, Lviv 79060, Ukraine, and

Institut für Informatik, Humboldt-Universität zu Berlin, Unter den Linden 6, Berlin 10099, Germany
verbitsky@informatik.hu-berlin.de

YAROSLAV VOROBETS received his Ph.D. from Moscow State University in 1998. Then he became a research scholar at Lviv University and the Institute for Applied Problems of Mechanics and Mathematics in Lviv, Ukraine. Since 2005 he has been with the Department of Mathematics at Texas A&M University. His research interests are focused on dynamical systems.

Department of Mathematics, Texas A&M University, College Station, TX 77843-3368, USA
yvorobet@math.tamu.edu

Zero

The tragedy of mankind.
The rational
Is a set of measure zero
In the real.

—Submitted by Rick Norwood, East Tennessee State University

© THE MATHEMATICAL ASSOCIATION OF AMERICA [Monthly 117]