A dynamical proof of Pisot’s theorem

Jaroslaw Kwapisz

Department of Mathematical Sciences
Montana State University
Bozeman MT 59717-2400
tel: (406) 994 5343
fax: (406) 994 1789
e-mail: jarek@math.montana.edu
web page: http://www.math.montana.edu/~jarek/

March 9, 2005

Abstract

We give a geometric proof of classical results that characterize Pisot numbers
as algebraic $\lambda > 1$ for which there is $x \neq 0$ with $\lambda^n x \to 0 \pmod{1}$ and identify
such $x$ as members of $\mathbb{Z}[\lambda^{-1}] : \mathbb{Z}[\lambda]^*$ where $\mathbb{Z}[\lambda]^*$ is the dual module of $\mathbb{Z}[\lambda]$.

A real number $\lambda > 1$ is called a Pisot number iff it is an algebraic integer and all
its Galois conjugates (other than $\lambda$) are of modulus less that one — the golden mean
$(1 + \sqrt{5})/2$ is an example. Pisot’s 1938 thesis [4] and, independently, Vijayaraghavan’s
1941 paper [7] contain the following beautiful characterization.

Theorem 1 (Pisot, Vijayaraghavan) Suppose that $\lambda > 1$ is an algebraic number
(over the field of rational numbers $\mathbb{Q}$). The following are equivalent

(i) $\lambda$ is a Pisot number;

(ii) There exists non-zero real $x$ such that $\lim_{n \to \infty} \lambda^n x = 0 \pmod{1}$ (i.e.
$\lim_{n \to \infty} \min \{|\lambda^n x - k| : k \in \mathbb{Z}\} = 0$ where $\mathbb{Z}$ are rational integers).

Moreover, any $x$ satisfying (ii) belongs to $\mathbb{Q}(\lambda)$, the field extension of $\mathbb{Q}$ by $\lambda$.

The property (ii) is responsible for Pisot numbers turning up in a variety of contexts
seemingly unrelated to their definition. The reader may want to savor the ensuing
connections by reading [5, 2]. Our interest in Pisot’s theorem stems from its role in
determination of spectrum for the translation flow on substitution tiling spaces, as
exhibited by [6] and further exploited in [1]. We shall not discuss that connection here.
and turn instead to our goal of supplying a proof of the theorem that offers a direct geometrical insight — something that is missing from the considerations of the classical proofs (as found in [3] or [5]). We shall also derive the following characterization of the set

\[ X_\lambda := \{ x \in \mathbb{R} : \lim_{n \to \infty} \lambda^n x = 0 \pmod{1} \}. \]  

(1)

In [3], this result is also attributed to Pisot and Vijayaraghavan.

**Theorem 2 (Pisot, Vijayaraghavan)** Suppose \( \lambda > 1 \) is Pisot. Let \( p' \) be the derivative of the monic irreducible polynomial of \( \lambda \) over \( \mathbb{Z} \), and \( \mathbb{Z}[\lambda]^* := \frac{1}{p'(\lambda)} \mathbb{Z}[\lambda] \). Then \( x \in X_\lambda \) if and only if \( \lambda^n x \in \mathbb{Z}[\lambda]^* \) for some \( n \geq 0 \); i.e.,

\[ X_\lambda = \bigcup_{n \geq 0} \lambda^{-n} \mathbb{Z}[\lambda]^* = \mathbb{Z}[\lambda^{-1}] \cdot \mathbb{Z}[\lambda]^*. \]  

(2)

We note that \( \mathbb{Z}[\lambda]^* \) is just an explicit form (as given by Euler) of the dual of the module \( \mathbb{Z}[\lambda] \) typically defined as \( \mathbb{Z}[\lambda]^* := \{ x \in \mathbb{Q}(\lambda) : \text{trace}(xy) \in \mathbb{Z} \forall y \in \mathbb{Z}[\lambda] \} \) and that \( \mathbb{Z}[\lambda]^* \) is non-zero only if \( \lambda \) is an algebraic integer (see Prop. 3-7-12 in [8]). That \( x \in X_\lambda \) for \( x \in \mathbb{Z}[\lambda]^* \) is clear by the following standard argument (emulating Theorem 1 in [5]). Let \( \lambda = \lambda_1, \lambda_2, \ldots, \lambda_d \) be all the roots of \( p \) (the Galois conjugates of \( \lambda \)) and \( x = x_1, \ldots, x_d \) be the images of \( x \) under the natural isomorphisms \( \mathbb{Q}(\lambda) \to \mathbb{Q}(\lambda_i), \)

\[ x_i \in \mathbb{Q}(\lambda_i). \]

Then

\[ Z \ni T_n := \text{trace}(\lambda^n x) = \sum_{i=1}^{d} \lambda_i^n x_i = \lambda^n x + \sum_{i=2}^{d} \lambda_i^n x_i, \]  

(3)

and so \( |\lambda^n x - T_n| \to 0 \) due to the Pisot hypothesis: \( |\lambda_i| < 1 \) for \( i = 2, \ldots, d \).

From now on, consider a fixed algebraic number \( \lambda > 1 \). Denote by \( p \) its monic minimal polynomial, which is obviously irreducible. Let \( d := \deg(p) \), and fix a \( d \times d \) matrix \( A \) over \( \mathbb{Q} \) with eigenvalue \( \lambda \). The companion matrix of \( p \) is one such \( A \), and any other is similar to it over \( \mathbb{Q} \). If \( \lambda \) is an algebraic integer then \( A \) can be taken over \( \mathbb{Z} \). Conversely, if \( A \) preserves some lattice in \( L \subset \mathbb{R}^d \), \( AL \subset L \), then \( \lambda \) is an algebraic integer. Here by a **lattice** we understand a discrete rank \( d \) subgroup of \( \mathbb{R}^d \) — \( \mathbb{Z}^d \) being the simplest example.

We shall frequently use the fact that \( A \) is **irreducible over** \( \mathbb{Q} \): if \( W \) is a non-zero subspace of \( \mathbb{Q}^d \) and \( A(W) \subset W \), then \( W = \mathbb{Q}^d \) (as otherwise the characteristic polynomial of \( A|_W \) would divide \( p \)). Also, by irreducibility of \( p \), \( A \) has simple eigenvalues and is diagonalizable over \( \mathbb{C} \) so that we have a splitting

\[ \mathbb{R}^d = E^s \oplus E^c \oplus E^u \]

where \( E^s, E^c, E^u \) are the linear spans of the real eigenspaces corresponding to the eigenvalues of modulus less, equal, and greater than 1, respectively. We shall see that, for \( v \in \mathbb{R}^d \setminus \{0\}, \) \( A^n v \to 0 \) iff \( v \in E^s \) lies at the very heart of Pisot’s theorem. Below, \( \langle \cdot | \cdot \rangle \) is the standard scalar product in \( \mathbb{R}^d \).
Lemma 1 If \( \langle A^n v_0 | k_0 \rangle \to 0 \) (mod 1) for some \( v_0 \in \mathbb{R}^d \setminus E^* \) and \( k_0 \in \mathbb{Z}^d \setminus \{0\} \), then 
A leaves invariant some lattice in \( \mathbb{Q}^d \); i.e., \( \lambda \) is an algebraic integer.

Lemma 2 Suppose that \( A \) has entries in \( \mathbb{Z} \) and \( k_0 \in \mathbb{Z}^d \setminus \{0\} \). If \( \langle A^n v_0 | k_0 \rangle \to 0 \) (mod 1) for \( v_0 \in \mathbb{R}^d \), then \( v_0 \in \mathbb{Q}^d + E^* \).

Proof of Theorem 1: Taking \( x = 1 \) in (3) shows that (i) implies (ii), so it is left to show (i) from (ii). Pick \( \omega \in \mathbb{R}^d \) to be an eigenvector of \( A \) corresponding to \( \lambda \), \( A \omega = \lambda \omega \).

Fix \( k_0 \in \mathbb{Z}^d \setminus \{0\} \). Observe that \( \langle k_0 | \omega \rangle \neq 0 \) by irreducibility of the transpose \( A^T \) of \( A \) (since \( \{ q \in \mathbb{Q}^d : \langle q | \omega \rangle = 0 \} \) is \( A^T \) invariant). Thus, in the linear span \( \text{lin}_{\mathbb{R}}(\omega) \) of \( \omega \) over \( \mathbb{R} \), we can find \( v_0 \) so that \( x = \langle v_0 | k_0 \rangle \). In this way,

\[
\lambda^n x = \lambda^n \langle v_0 | k_0 \rangle = \langle A^n v_0 | k_0 \rangle, \quad v_0 \in \text{lin}_{\mathbb{R}}(\omega). 
\] (4)

From \( x \neq 0 \), \( v_0 \not\in E^* \) and so \( \lambda \) must be an algebraic integer by Lemma 1. By Lemma 2, \( v_0 = q_0 + z \) for some \( z \in E^* \) and \( q_0 \in \mathbb{Q}^d \), and \( q_0 \neq 0 \) from \( v_0 \not\in E^* \). Consider, \( W := \mathbb{Q}^d \cap (E^* \oplus \text{lin}_{\mathbb{R}}(\omega)) \). Irreducibility of \( A \), \( AW \subset W \) and \( q_0 \in W \) force \( W = \mathbb{Q}^d \). Thus \( E^* \oplus \text{lin}_{\mathbb{R}}(\omega) = \mathbb{R}^d \) and \( \lambda \) is Pisot. □

We turn our attention to proving the lemmas now. The two proofs will partially overlap and could be combined into a single more compact argument, but we shall keep them separate because (in applications) \( \lambda \) is often a priori known to be an algebraic integer. In that case, Pisot’s theorem can be viewed as a feature of the dynamics of the endomorphism \( f : \mathbb{T}^d \to \mathbb{T}^d, x \pmod{\mathbb{Z}^d} \mapsto Ax \pmod{\mathbb{Z}^d} \), induced by \( A \) on the \( d \)-dimensional torus, \( \mathbb{T}^d := \mathbb{R}^d / \mathbb{Z}^d \). Beside the toral endomorphism \( f \), our main tool will be the concept of duality of lattices. Recall that the dual of a lattice \( L \) is defined as \( L^* := \{ v \in \mathbb{R}^d : \langle v | l \rangle \in \mathbb{Z} \ \forall l \in L \} \). One easily checks that \( (\mathbb{Z}^d)^* = \mathbb{Z}^d \). For any lattice \( L \), after expressing it as \( L = B \mathbb{Z}^d \) for some nonsingular matrix \( B \), we have \( L^* = (B \mathbb{Z}^d)^* = (B^T)^{-1} \mathbb{Z}^d \) where \( B^T \) is the transpose of \( B \). In particular, \( L^* \) is also a lattice.

Proof of Lemma 1: Let \( V := \{ v \in \mathbb{R}^d : \langle A^n v | k_0 \rangle \to 0 \) (mod 1) \} and \( K := \{ k \in \mathbb{Q}^d : \langle A^n v | k \rangle \to 0 \) (mod 1) \} \} for all \( v \in V \}. These are subgroups of \( \mathbb{R}^d \), \( A(V) = V \), \( A(T(K) = K \), and \( v_0 \in V, k_0 \in K \). Irreducibility of \( A^T \) forces \( \text{lin}_{\mathbb{Q}}(K) = \mathbb{Q}^d \) so that we can find linearly independent \( k_1, \ldots, k_d \in K \). Let \( \Gamma \) be the lattice generated by \( k_j \)'s, \( \Gamma^* \) be its dual, and \( \chi_j : \mathbb{R}^d / \Gamma^* \to \mathbb{C} \) be the associated basis characters on the torus \( \mathbb{R}^d / \Gamma^* \); namely, \( \chi_j(x \pmod{\Gamma^*}) := \exp(2\pi i (k_j \cdot x)), x \in \mathbb{R}^d, j = 1, \ldots, d \).

The convergence \( \langle A^n v_0 | k_j \rangle \to 0 \) (mod 1) translates to \( \chi_j(A^n v_0 \pmod{\Gamma^*}) \to 1 \), which (by continuity of \( \chi_j \) and compactness of \( \mathbb{R}^d / \Gamma^* \)) is equivalent to \( \text{dist}(A^n v_0 \pmod{\Gamma^*}, \chi_j^{-1}(1)) \to 0 \). Therefore, \( \text{dist}(A^n v_0 \pmod{\Gamma^*}, G) \to 0 \) where \( G := \bigcap_{j=1}^d \chi_j^{-1}(1) = \{ 0 \pmod{\Gamma^*} \} \), which is to say that

\[
\text{dist}(A^n v_0, \Gamma^*) \to 0. 
\] (5)

Fix \( \epsilon > 0 \) so that, for \( x, y \in \Gamma^* \cup \Gamma^* \), \( \text{dist}(x, y) < \epsilon \) forces \( x = y \). (This is possible because \( \Gamma^* / \Gamma^* \) is discrete in \( \mathbb{R}^d / \Gamma^* \), as can be seen by picking \( a \in \mathbb{N} \) so that \( aA \) has all integer entries and observing that \( \mathbb{A} \mathbb{G}^* \subset a^{-1} \Gamma^* \), which yields \( \Gamma^* / \Gamma^* \subset (a^{-1} \Gamma^*) / \Gamma^* \).)
From (5), there are \( u_n \in \Gamma^* \), \( n \in \mathbb{N} \), such that \( \text{dist}(A^n v_0, u_n) \to 0 \). Since, \( \text{dist}(u_{n+1}, A u_n) \leq \text{dist}(u_{n+1}, A^{n+1} v_0) + \text{dist}(A A^n v_0, A u_n) \), we have \( \text{dist}(u_{n+1}, A u_n) \to 0 \) and so, as soon as \( \text{dist}(u_{n+1}, A u_n) < \epsilon \), it must be that \( u_{n+1} = A u_n \). Therefore, for some \( n_0 \in \mathbb{N} \) and all \( l \geq 0 \), we have \( A^l u_{n_0} = u_{n_0 + l} \in \Gamma^* \). Now, from \( v_0 \not\in E^* \), \( A^n v_0 \not\in \Gamma^* \) so that \( u_{n_0} \neq 0 \). But \( u_{n_0} \in M := \{ v \in \Gamma^* : A^l v \in \Gamma^* \forall l \geq 0 \} \), which makes \( M \) a nonzero subgroup of \( \Gamma^* \). Clearly \( A M \subset M \). By irreducibility of \( A \), \( \text{lin}_Q(M) = Q^d \) so that \( M \) is a lattice. \( \square \)

**Proof of Lemma 2:** Let \( f : \mathbb{T}^d \to \mathbb{T}^d \) be the toral endomorphism associated to \( A \), \( \chi : \mathbb{T}^d \to \mathbb{C} \) be the character associated to \( k_0 \), \( \chi(x (\text{mod} \mathbb{Z}^d)) := \exp(2\pi i (x | k_0)) \), and set \( p := v_0 \) (mod \( \mathbb{Z}^d \)). The hypothesis \( \langle A^n v_0 | k_0 \rangle \to 0 \) (mod 1) translates to \( \chi(f^n(p)) \to 1 \), which is equivalent to \( \text{dist}(f^n(p), G) \to 0 \) where \( G := \chi^{-1}(1) \). We claim that, in fact,

\[
\text{dist}(f^n(p), G_\infty) \to 0, \quad G_\infty := \bigcap_{n \geq 0} f^{-n}(G). \quad (6)
\]

Indeed, otherwise \( f^{n_k}(p) \to w \not\in f^{-l}(G) \) for some \( w \), \( l \geq 0 \), and \( n_k \to \infty \); and so \( f^{n_k+l}(p) \to f^l(w) \not\in G \) contradicting \( \text{dist}(f^n(p), G) \to 0 \).

To identify \( G_\infty \) as a finite subgroup of \( \mathbb{T}^d \), consider its lift to \( \mathbb{R}^d \),

\[
\Gamma := G_\infty + \mathbb{Z}^d := \{ x \in \mathbb{R}^d : x (\text{mod} \mathbb{Z}^d) \in G_\infty \}.
\]

Denote by \( L_{k_0} \) the smallest sublattice of \( \mathbb{Z}^d \) containing \( (A^T)^n k_0 \) for all \( n \geq 0 \). Its dual, \( L_{k_0}^* \), is a lattice in \( \mathbb{Q}^d \). For \( v \in \mathbb{R}^d \), we have \( v \in \Gamma \) iff \( \langle A^n v | k_0 \rangle = \langle v | (A^T)^n k_0 \rangle \in \mathbb{Z} \) for all \( n \geq 0 \) iff \( v \in L_{k_0}^* \). Thus \( G_\infty = \Gamma / \mathbb{Z}^d \) where

\[
\Gamma = L_{k_0}^* \subset \mathbb{Q}^d. \quad (7)
\]

Let \( q_n \in G_\infty \) realize the distance in (6) so that \( \text{dist}(f^n(p), q_n) \to 0 \) and thus also \( \text{dist}(f(q_n), q_{n+1}) \to 0 \). Since \( G_\infty \) is discrete, there is \( n_0 \in \mathbb{N} \) such that

\[
q_{n+1} = f(q_n), \quad n \geq n_0. \quad (8)
\]

Moreover, if we pick \( \epsilon > 0 \) small enough and \( n_1 > n_0 \) large enough, then for every \( n \geq n_1 \) we can write \( f^n(p) = q_n + x_n + y_n + z_n \) for some unique \( x_n \in E^s \), \( y_n \in E^c \), \( z_n \in E^u \), each of norm less than \( \epsilon \). From (8), we have \( x_{n+1} = Ax_n \), \( y_{n+1} = Ay_n \), \( z_{n+1} = Az_n \) for \( n \geq n_1 \). What is more, \( \text{dist}(f^n(p), q_n) \to 0 \) forces \( y_n \to 0 \) and \( z_n \to 0 \), which is only possible if \( y_{n_1} = 0 \) and \( z_{n_1} = 0 \). Thus \( f^{n_1}(p) = q_{n_1} + x_{n_1} \); i.e., \( A^{n_1} v_0 = w + x_{n_1} \) for some \( w \in \Gamma \) (with \( q_{n_1} = w \) (mod \( \mathbb{Z}^d \))). To summarize, \( v_0 \in A^{-n_1} \Gamma + E^s = A^{-n_1} L_{k_0}^* + E^s \subset Q^d + E^s \). \( \square \)

**Remark 1 (addendum to Lemma 2)** Under the hypotheses of Lemma 2,

\[
\{ v \in \mathbb{R}^d : \langle A^n v | k_0 \rangle \to 0 \text{ (mod 1)} \} = \bigcup_{n \geq 0} A^{-n} L_{k_0}^* + E^s \quad (9)
\]

where \( L_{k_0} \) is the smallest lattice in \( \mathbb{Z}^d \) containing \( (A^T)^n k_0 \) for all \( n \geq 0 \).
Proof of Remark 1: The “⊂” inclusion is demonstrated in the proof of Lemma 2. To see “⊃”, it suffices to note that, if \( v \in L_{k_0}^* + E^s \), then \( v = w + x \) where \( w \pmod{Z^d} \in G_\infty \) and \( x \in E^s \). Thus \( \langle A^n v | k_0 \rangle \) becomes exponentially close to \( \langle A^n w | k_0 \rangle \in \mathbb{Z} \) as \( n \to \infty \). □

Proof of Theorem 2: The plan is to explicitly compute the objects involved in the preceding arguments for \( A \) that is the companion matrix of the polynomial \( p \) of \( \lambda \),

\[
p(x) = x^d + a_{d-1}x^{d-1} + \cdots + a_0, \quad a_i \in \mathbb{Z}.
\]

The eigenvectors \( \omega \) and \( \omega^* \) with \( A\omega = \lambda \omega \), \( A^T\omega^* = \lambda \omega^* \) can be found as

\[
\omega^* := \frac{1}{p'(\lambda)} \cdot (a_1 + a_2\lambda + \cdots + \lambda^{d-1}, a_{d-1} + \lambda, 1)
\]
\[
\omega := (1, \lambda, \lambda^2, \ldots, \lambda^{d-1}).
\]

These are normalized so that \( \langle \omega | \omega^* \rangle = 1 \), which ensures that the projection onto \( \text{lin}_\mathbb{R}(\omega) \) along \( E^s = (\omega^*)^\perp \) is given by \( \text{pr}^u(y) = \langle y | \omega^* \rangle \omega \), \( y \in \mathbb{R}^d \). Note that the components of \( \omega^* \) generate \( \frac{1}{p'(\lambda)} \mathbb{Z}[\lambda] \), \( \{ \langle u | \omega^* \rangle : u \in \mathbb{Z}^d \} = \frac{1}{p'(\lambda)} \mathbb{Z}[\lambda] \).

Denote by \( e_1, \ldots, e_d \) the standard basis in \( \mathbb{R}^d \), and set \( k_0 := e_1 \). Since \( e_i = (A^T)^{-1}(e_1) \) for \( i = 1, \ldots, d \), we have \( L_{k_0} = \mathbb{Z}^d \). Hence, \( L_{k_0}^* = \mathbb{Z}^d \).

If we write \( x = \langle v_0 | k_0 \rangle \) for \( v_0 \in \text{lin}_\mathbb{R}(\omega) \) — as in (4) in the proof of Theorem 1 — then \( \lambda^nx \to 0 \pmod{1} \) iff \( \langle A^n v_0 | k_0 \rangle \to 0 \pmod{1} \) iff \( A^{n_1}v_0 \in L_{k_0}^* + E^s = \mathbb{Z}^d + E^s \) for some \( n_1 \geq 0 \), where the last equivalence hinges on Remark 1. Thus \( x \in X_\lambda \) are of the form

\[
x = \lambda^{-n_1} \langle A^{n_1}v_0 | k_0 \rangle = \lambda^{-n_1} \langle \text{pr}^u(u) | k_0 \rangle = \lambda^{-n_1} \langle u | \omega^* \rangle \langle \omega | k_0 \rangle = \lambda^{-n_1} \langle u | \omega^* \rangle \cdot 1 \quad (10)
\]

where \( u \in \mathbb{Z}^d \) and \( n_1 \geq 0 \). That is \( X_\lambda = \bigcup_{n_1 \geq 0} \lambda^{-n_1} \frac{1}{p'(\lambda)} \mathbb{Z}[\lambda] \), as desired. □

The readers accustomed to a more traditional framework will no doubt notice that, in our setting, the scalar product \( \langle \cdot | \cdot \rangle \) on \( \mathbb{R}^d \times \mathbb{R}^d \) serves as the completion of the trace form on \( \mathbb{Q}(\lambda) \times \mathbb{Q}(\lambda) \), the two being related by \( \langle x | y \rangle = \text{trace}(\langle x | \omega^* \rangle \cdot \langle \omega | y \rangle) \) for \( x, y \in \mathbb{Q}^d \). This explains our remark about the nature of \( \mathbb{Z}[\lambda]^* \) from the beginning of this note.

References


