Arithmetic Dynamics

The study of the number-theoretic properties of rational and algebraic points under repeated application of a polynomial or rational function.

Shihan Kanungo

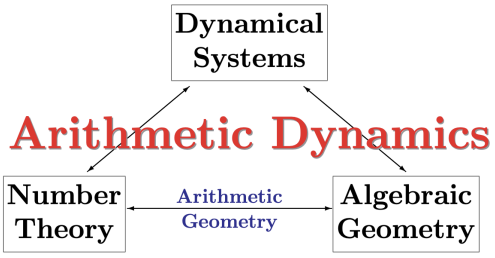
Euler Circle, Palo Alto, CA 94306

Abstract

In this expository paper, we investigate the topic of *arithmetic dynamics*. Discrete dynamical systems is the study of iteration of functions. Number theory is the study of properties of integers. Combine the two and you land in the brave new world of arithmetic dynamics, wherein we study number theoretic properties of orbits of integers and rational numbers under iteration of polynomials and rational functions. The main idea of arithmetic dynamics is that we take a function from some set to itself, and we look at how it behaves as we iterate it over and over. In particular, we will first define the set of p-adic numbers and introduce some useful results regarding them. Then we will look at an application of arithmetic dynamics, and we will relate it to dynamics in the p-adic numbers.

1 Introduction

In this paper, we investigate the topic of *arithmetic dynamics*. Whereas classical discrete dynamics is the study of iteration of self-maps of the complex plane or real line, arithmetic dynamics is the study of the number-theoretic properties of rational and algebraic points under repeated application of a polynomial or rational function.



- **Number Theory:** Study properties of integers and rational numbers
- **Algebraic Geometry:** Study solutions to systems of polynomial equations
- **(Discrete) Dynamical Systems:** Study orbits of iteration of functions
- **Arithmetic Dynamics:** Study number-theoretic properties of orbits of rational numbers for iteration of polynomial functions.

Figure courtesy of Joseph H. Silverman, *Arithmetic Dynamics: A Survey*, International Congress of Mathematicians 2022

A principal theme of arithmetic dynamics is that many of the fundamental problems in the theory of Diophantine equations have dynamical analogs.

While number theory looks for patterns in sequences of numbers, dynamical systems actually produce sequences of numbers — like the sequence that defines a planet's position in space at regular intervals of time. The two merge when mathematicians look for number–theoretic patterns hidden in those sequences.

Global arithmetic dynamics is the study of analogues of classical Diophantine geometry in the setting of discrete dynamical systems, while local arithmetic dynamics, also called p-adic or non-archimedean dynamics, is an analogue of complex dynamics in which one replaces the complex numbers $\mathbb{C}$ by a p-adic field such as $\mathbb{Q}_p$ or $\mathbb{C}_p$ and studies chaotic behavior and the Fatou and Julia sets.

2 *p*-adic Numbers

We will now define the set of p-adic numbers and introduce the basic theory. To start off, we need to define magnitudes and distances on the rationals using the multiplicity of primes in the factorization of numbers. This is done by the absolute value function.

The *p*-adic valuation

Prior to the absolute value itself, we will construct a valuation, which is a function that relates each integer to the multiplicity of a prime in its factorization.

Definition 1. A function $v : \mathbb{Z} \to \mathbb{Z} \cup \{\infty\}$ is called a *valuation* if for all $x, y \in \mathbb{Z}$, we have

i. $v(x, y) = v(x) + v(y)$

ii. $v(x + y) \geqslant \min(v(x), v(y))$

iii. $v(0) = \infty$

Then we define the p-adic valuation:

Definition 2. Let p be a prime number. For $x \in \mathbb{Z}\backslash\{0\}$, we define $v_p(x)$ to be the unique nonnegative integer k such that

$$x = p^k x',$$

where $x' \in \mathbb{Z}$ and $p \nmid x$. Also define $v_p(0) = \infty$.

We can also think of $v_p(x)$ as the exponent of p in the prime factorization of x.

Example 3.

– If $p = 3$ and $x = 54$, then $x = 2 \cdot 3^3$, so $v_p(x) = 3$.

– If $p = 5$ and $x = 42$, then $5 \nmid x$, so $v_p(x) = 0$.

– If $p = 2$ and $x = 40$, then $x = 2^3 \cdot 5$, so $v_p(x) = 3$.

We can extend the p-adic valuation to the rational numbers as follows: for $x \in \mathbb{Q}$ where $x = a/b$, define

$$v_p(x) = v_p(a) - v_p(b).$$

Example 4.

– We have $v_3\left(\frac{36}{41}\right) = v_3(36) - v_3(41) = 2 - 0 = 0$.

– We have $v_7\left(\frac{11}{98}\right) = v_7(11) - v_7(98) = 0 - 2 = -2$.

– We have $v_3\left(\frac{3}{27}\right) = v_3(3) - v_3(27) = 1 - 3 = -2$.

From these examples, we can see that the extension of the function to all the rationals is well-defined for equivalent fractions, and that representing integers as rational numbers does not change the value of the function. In general, if $q = \frac{a}{b}$ is a rational number, $v_p\left(\frac{a}{b}\right)$ does not depend on the choice of a and b.

We are now ready to define the p-adic absolute value.

Definition 5. If $x \in \mathbb{Q}$, we define $|x|_p = p^{-v_p(x)}$ if $x \neq 0$ and $|0|_p = 0$.

Example 6.

– We have $|6|_3 = 3^{-v_3(6)} = 3^{-1} = \frac{1}{3}$.

– We have $|32|_7 = 7^{-v_7(32)} = 7^0 = 1$.

– We have $\left|\frac{1}{40}\right|_5 = 5^{v_p\left(\frac{1}{40}\right)} = 5^{-(-1)} = 5$.

The p-adic absolute value shares some important properties with the standard absolute value.

Theorem 7. *For all $x, y \in \mathbb{Q}$, we have*

i. $|x|_p = 0 \Leftrightarrow x = 0$

ii. $|xy|_p = |x|_p \cdot |y|_p$

iii. $|x + y|_p \leqslant \max(|x|_p, |y|_p)$

iv. $|x + y|_p \leqslant |x|_p + |y|_p$ (Triangle Inequality)

Remark. An *absolute value* on a field F is a map $|\cdot| : F \to \mathbb{R}^+$ that satisfies properties i, ii, and iv. An absolute value that also satisfies property iii, is said to be *non-archimedean*.

Proof. Let p be a prime number, and let $x \in \mathbb{Q}$.

i. If $x \neq 0$, we have $|x|_p = p^{-v_p(x)}$, which can never be zero.

ii. Write $x = p^{v_p(x)}x'$ and $y = p^{v_p(y)}y'$ so $v_p(x') = v_p(y') = 0$. Then

$$\begin{aligned} |xy|_p &= \left|p^{v_p(x)}x'p^{v_p(y)}y'\right|_p \\ &= \left|p^{v_p(x)+v_p(y)}x'y'\right|_p \\ &= p^{-(v_p(x)+v_p(y))} \\ &= p^{-v_p(x)}p^{-v_p(y)} \\ &= |x|_p|y|_p. \end{aligned}$$

iii. Without loss of generality, take $|x|_p \geqslant |y|_p$. Then, we have that $p^{-v_p(x)} \geqslant p^{-v_p(y)}$, which means $v_p(x) \leqslant v_p(y)$. Therefore, $\min\{v_p(x), v_p(y)\} = v_p(x)$. It is easily verified that

$$v_p(x + y) \geqslant \min\{v_p(x), v_p(y)\} = v_p(x).$$

Therefore, $p^{-v_p(x+y)} \leqslant p^{-v_p(x)}$, or equivalently, $|x + y|_p \leqslant |x|_p = \max\{|x|_p, |y|_p\}$.

iv. This follows directly from iii. as we have

$$|x + y|_p \leqslant \max\left\{|x|_p, |y|_p\right\} \leqslant |x|_p + |y|_p$$

for all $x, y \in \mathbb{Q}$. □

A convergent sequence (x_m) is characterized by the fact that its terms become (and stay) arbitrarily close to its limit, as $m \to \infty$. Due to this, however, they also get close to each other; in fact, $|x_m - x_n|$ can be made arbitrarily small for sufficiently large m and n. It is

natural to ask whether the latter property, in turn, implies the existence of a limit. This problem was first studied by Augustin-Louis Cauchy (1789–1857). Thus we shall call such sequences *Cauchy sequences*. More precisely, we have the following.

Definition 8. Given an absolute value, $|\cdot|$, on F, we say a sequence (x_k) in F is *Cauchy* if for every positive real number $\varepsilon > 0$ there is a positive integer N such that for all natural numbers $m, n > N$,

$$|x_m - x_n| < \varepsilon.$$

We say F is *complete* with respect to $|\cdot|$ if every Cauchy sequence in F has a limit in F. Further, we denote by $\mathfrak{C}_p$ the set of all sequences in $\mathbb{Q}$ which are Cauchy with respect to $|\cdot|_p$.

Example 9.

(1) The sequence (a_n), where $a_n = 1 - 2^{-n}$,

$$1 + \tfrac{1}{2},\ 1 - \tfrac{1}{4},\ 1 + \tfrac{1}{8},\ 1 - \tfrac{1}{16},\ \dots$$

is Cauchy with respect to the standard absolute value. It converges to 1.

(2) Let p be a prime. The sequence (b_n), where $b_n = p + \cdots + p^n$,

$$p,\ p + p^2,\ p + p^2 + p^3,\ \dots$$

is Cauchy in $\mathbb{Q}$ with respect to $|\cdot|_p$. It does not have a limit in $\mathbb{Q}$, but we will see later that it converges to a p-adic number.

We can add and multiply Cauchy sequences. The constant sequences $0 = (0, 0, \dots)$ and $1 = (1, 1, \dots)$ are additive and multiplicative identities, and every Cauchy sequence (x_n) has an additive inverse $(-x_n)$. So Cauchy sequences form a commutative ring.

Proposition 10. *Under the operations* $(x_k) + (y_k) = (x_k + y_k)$ *and* $(x_k) \cdot (y_k) = (x_k y_k)$*, the set* $\mathfrak{C}_p$ *of Cauchy sequences in* $\mathbb{Q}$ *is a commutative ring with identity.*

Proof. All we need to check is that if (x_k) and (y_k) are Cauchy, then $(x_k + y_k)$ and $(x_k y_k)$ are Cauchy. Suppose $\varepsilon > 0$. Since x and y are Cauchy, there exists $N > 0$ such that for all $m, n > N$, we have that $|x_m - x_n|_p < \frac{\varepsilon}{2}$ and $|y_m - y_n|_p < \frac{\varepsilon}{2}$ (just find N separately for (x_k) and (y_k) and take the larger value). Then, by the triangle inequality, we have

$$\begin{aligned} |(x_m + y_m) - (x_n + y_n)|_p &= |x_m - x_n + y_m - y_n|_p \\ &\leqslant x_m - x_n|_p + |y_m - y_n|_p \\ &< \varepsilon. \end{aligned}$$

Thus $(x_k + y_k)$ is Cauchy. Now we show $(x_n)(y_n)$ is Cauchy. Let $\varepsilon > 0$. Then

$$\exists\ B\ :\ |x_n|_p, |y_n|_p < \tfrac{B}{2}\ \text{ for all } n.$$

Also,

$$\exists\ N\ :\ m, n > N\ \Rightarrow\ |x_m - x_n|_p, |y_m - y_n|_p < \tfrac{\varepsilon}{B}$$

Then for all $m, n > N$, we have

$$\begin{aligned}
|x_m y_m - x_n y_n|_p &= |x_m y_m - x_m y_n + x_m y_n - x_n y_n|_p \\
&= |x_m(y_m - y_n) + y_n(x_m - x_n)|_p \\
&\leqslant |x_m(y_m - y_n)|_p + |y_n(x_m - x_n)|_p \\
&\leqslant |x_m|_p \cdot |y_m - y_n|_p + |y_n|_p \cdot |x_m - x_n|_p \\
&< \frac{B}{2} \cdot \frac{\varepsilon}{B} + \frac{B}{2} \cdot \frac{\varepsilon}{B} \\
&= \varepsilon.
\end{aligned}$$

Thus $(x_n)(y_n)$ is Cauchy. □

The field of p-adic numbers

An *ideal* I of a commutative ring R is a subset of R which is a group under addition, and also has the property that it is closed under multiplication by any $r \in R$, that is, for all $r \in R$ and $x \in I$, $rx \in I$. An ideal of the form

$$(x) = xR = \{xr \mid r \in R\}$$

is called a *principal ideal*. Note that the commutative ring R itself is a principal ideal, $R = (1)$. An ideal that is contained in no other ideal except R is called a *maximal ideal*.

Example 11. The set of even numbers E is an ideal of the integers. It is in fact a maximal ideal because it we add an odd number $2n + 1$ to it, we can subtract by $2n \in E$ to get 1.

We can define the *sum* of two ideals as the Minkowsky sum, i.e.

$$I_1 + I_2 = \{x_1 + x_2 \mid x_1 \in I_1, x_2 \in I_2\}.$$

We write (x, y) to denote $xR + yR$. Let R be a commutative ring and let I be an ideal in R. We can form the factor group R/I whose elements are the sets

$$x + I = \{x\} + I = \{x + y \mid y \in I\}.$$

We define addition and multiplication of elements of R/I by

$$\begin{aligned}
(x + I) + (y + I) &= (x + y) + I, \\
(x + I) \cdot (y + I) &= (xy) + I.
\end{aligned}$$

We then have the following well-known theorem:

Theorem 12. *Let R be a commutative ring with identity, and let I be an ideal of R. Then R/I is a field if and only if I is a maximal ideal.*

Denote by $\mathfrak{N}$ the set of sequences (x_k) in $\mathfrak{C}_p$ such that $\lim_{k\to\infty} |x_k|_p = 0$.

Theorem 13. *The ideal $\mathfrak{N}$ is a maximal ideal in $\mathfrak{C}_p$.*

Proof. Let $(x_n) \in \mathfrak{C}_p \backslash \mathfrak{N}$, and let $\mathfrak{I} = ((x_k), \mathfrak{N})$, i.e. $\mathfrak{I}$ is the smallest ideal containing both (x_k) and N. We will show that $\mathfrak{I} = \mathfrak{C}_\mathfrak{p}$.

Since $\lim_{k\to\infty} |x_k|_p \neq 0$, there exists $c > 0$ and N such that

$$|x_k|_p > c > 0 \quad \text{for all} \quad k > N.$$

Thus for all $k > N$ we have that $x_k \neq 0$, so we can define the sequence (y_k) by

$$y_k = \begin{cases} 0 & \text{if } k < N \\ \frac{1}{x_l} & \text{if } k \geqslant N. \end{cases}$$

Now we show that (y_k) is Cauchy. We have

$$|y_{k+1} - y_k|_p = \left| \frac{1}{x_{k+1}} - \frac{1}{x_k} \right|_p = \left| \frac{x_k - x_{k+1}}{x_k x_{k+1}} \right|_p \leqslant \frac{1}{c^2} \cdot |x_k - x_{k+1}|_p,$$

for $k \geqslant N$, which goes to 0 as k goes to ∞. Because $|\cdot|_p$ is non-archimedian, we have that

$$\begin{aligned} |y_j - y_k|_p &= |y_j - y_{j-1} + y_{j-1} - y_{j-2} + \cdots + y_{k+1} - y_k|_p \\ &\leqslant \max |y_j - y_{j-1}|_p, |y_{j-1} - y_{j-2}|_p, \ldots |y_{k+1} - y_k|_p, \end{aligned}$$

which goes to 0 as $j, k \to \infty$. Thus (y_k) is Cauchy. Then

$$x_k y_k = \begin{cases} 0 & \text{if } k < N \\ 1 & \text{if } k \geqslant N, \end{cases}$$

so

$$(1) - (x_k)(y_k) = \begin{cases} 1 & \text{if } k < N \\ 0 & \text{if } k \geqslant N \end{cases} \quad \in \mathfrak{N}.$$

Thus (1) is the sum of an element of $\mathfrak{N} \subset \mathfrak{I}$ plus the product of an element of $\mathfrak{I}$ times an element of $\mathfrak{C}_p$, which is an element of $\mathfrak{I}$, so $(1) \in \mathfrak{I}$. Thus $\mathfrak{I} = \mathfrak{C}_p$, and the proof is complete. □

The quotient $\mathfrak{C}_p/\mathfrak{N}$ is thus a commutative ring with identity. We define

$$\mathbb{Q}_p = \mathfrak{C}_p/\mathfrak{N}.$$

We call $\mathbb{Q}_p$ the *field of p-adic numbers*.

Note that the proof of Theorem 5 also shows the following, which we record as a lemma.

Lemma 14. *A rational sequence (x_k) is Cauchy with respect to $|\cdot|_p$ if and only if*

$$|x_{k+1} - x_k|_p \to 0.$$

Thus $\mathbb{Q}_p$ is the set of equivalence classes of rational sequences that are Cauchy with respect to $|\cdot|_p$, where two sequences are equivalent when their difference converges to 0. Consider the rational sequence

$$x_n = \sum_{k=n_0}^{n} d_k p^k, \qquad d_k \in \{0, 1, \ldots, p-1\}, \quad d_{n_0} \neq 0, \tag{2.1}$$

where n_0 is an integer. This is a Cauchy sequence, so it makes sense to talk about its limit

$$x = \sum_{k=n_0}^{\infty} d_k p^k. \tag{2.2}$$

We define $|x|_p$ to be p^{-n_0}. We want to show that every equivalence class contains a sequence of the type given in (2.1) (an infinite expansion in base p).

We start with a lemma.

Lemma 15. *Let $(y_k) \in \mathfrak{C}_p \backslash \mathfrak{N}$. Then the sequence $(|y_k|_p)$ is eventually constant.*

Proof. Since $(y_k) \notin \mathfrak{N}$, there exists $c > 0$ and N_1 such that

$$n \geqslant N_1 \quad \Rightarrow \quad |y_n|_p > c$$

Since $(y_k) \in \mathfrak{C}_p$ there exists N_2 such that

$$n \geqslant N_2 \quad \Rightarrow \quad |y_{n+1} - y_n|_p < c$$

Let $N = \max(N_1, N_2)$. For all $n \geqslant N$, we have that

$$|y_n|_p = |y_n - y_{n+1} + y_{n+1} \leqslant \max(|y_n - y_{n+1}|_p, |y_{n+1}|_p) = |y_{n+1}|_p$$
$$|y_{n+1}|_p = |y_{n+1} - y_n + y_n \leqslant \max(|y_{n+1} - y_n|_p, |y_n|_p) = |y_n|_p,$$

so $|y_n|_p = |y_{n+1}|_p$, which completes the proof. □

Then we have the following proposition:

Proposition 16. *Every equivalence class of $\mathfrak{C}_p$ contains a sequence of the type*

$$(\dagger) \qquad x_n = \sum_{k=n_0}^{n} d_k p^k, \qquad d_k \in \{0, 1, \ldots, p-1\}, \quad d_{n_0} \neq 0,$$

where n_0 is an integer.

Proof. Suppose $(y_k) \in \mathfrak{C}_p$. If $(y_k) \in \mathfrak{N}$, then the equivalence class of (y_k) contains (0), which is of the type (†). Otherwise, let p^{-n_0} be the value that $|y_k|_p$ takes as $k \to \infty$. Without loss of generality, we can replace (y_k) with another sequence such that $|y_k| = p^{-n_0}$ for all k. Define $y'_n = y_n p^{-n_0}$, so that $|y'_n| = 1$. Choose a subsequence (z_n) of (y'_n) such that

$$|z_{n+1} - z_n|_p \leqslant p^{n-1} \quad \text{for all } n.$$

Since $|z_n|_p = 1$, we can find $d'_0, d'_1, \ldots, d'_n \in \{0, 1, \ldots, p-1\}$ such that

$$x'_n = \sum_{k=0}^{n} d'_k p^n$$

has the property that $z_n \equiv x'_n \pmod{p^n}$. Since $|z_{n+1} - z_n|_p \leqslant p^{n-1}$, we have

$$z_{n+1} \equiv z_n \pmod{p^n},$$

so we can find d'_{n+1} such that $z_{n+1} \equiv x'_{n+1} \pmod{p^{n+1}}$. Continue like this to get a sequence (x'_n) of the type (†) that is equivalent to (y_n) and (z_n). □

The p-adic integers

Thus, we can think of $\mathbb{Q}_p$ as the set of base p expansions that can extend infinitely to the left and have finitely many digits to the right of the decimal point.
Now we define

$$\mathbb{Z}_p = \{x \in \mathbb{Q}_p \mid |x|_p \leqslant 1\}.$$

We call $\mathbb{Z}_p$ the set of p-adic integers. Then we can think of $\mathbb{Z}_p$ as the closed unit disk in $\mathbb{Q}_p$. Then $\mathbb{Z}_p$ is also a commutative ring with identity. We can think of the p-adic integers as infinite expansions (to the left) in base p that have no decimal places.

Example 17. The sequence in part (2) of Example 8 converges to the p-adic number given by $n_0 = 1$ and $d_k = 1$ for all $k \geqslant 1$. We can think of this p-adic number as the infinite expansion in base p: $\overline{\cdots 111}_p$. This number is also a p-adic integer.

Consider the set

$$p\mathbb{Z}_p = \{x \in \mathbb{Q}_p \mid |x|_p < 1\}.$$

This is an ideal of $\mathbb{Z}_p$. Consider the set

$$\mathbb{Z}_p \backslash p\mathbb{Z}_p = \{x \in \mathbb{Q}_p \mid |x|_p = 1\}.$$

This consists of all the invertible elements of $\mathbb{Z}_p$. Then any ideal that properly contains $p\mathbb{Z}_p$ contains an invertible element, which means it contains 1, so it is $\mathbb{Z}_p$. Thus $p\mathbb{Z}_p$ is a maximal ideal.

Lemma 18. *The field $\mathbb{Z}_p/p\mathbb{Z}_p$ has p elements.*

Proof. The elements of $\mathbb{Z}_p$ correspond to the expressions (2.2) with $n_0 \geqslant 0$. Two such sequences are equivalent iff they only differ in the first digit d_0. This gives p different elements of $\mathbb{Z}_p/p\mathbb{Z}_p$, corresponding to the different values of d_0. □

Then we have *Hensel's Lemma*, a useful lemma we use in a later part of the paper.

Lemma 19 (Hensel's Lemma)**.** *Let $f(x)$ be a polynomial in $\mathbb{Z}_p[x]$. If there exist $\alpha_1 \in \mathbb{Z}_p$ such that*

$$f(\alpha_1) \equiv 0 \pmod{p} \quad \textit{and} \quad f'(\alpha_1) \not\equiv 0 \pmod{p}$$

then there exists a unique p-adic integer α such that

$$f(\alpha) = 0$$

and $\alpha \equiv \alpha_1 \pmod{p}$.

Proof. We will construct a Cauchy sequence of p-adic integers $\alpha_1, \alpha_2, \ldots$ such that for all $n \geqslant 1$ we have that

$$f(\alpha_n) \equiv 0 \pmod{p^n} \quad \text{and} \quad \alpha_{n+1} \equiv \alpha_n \pmod{p^n}.$$

This sequence is Cauchy because $|\alpha_{n+1} - \alpha_n|_p \leqslant p^{-n}$. If α is the limit of this sequence, then note that $f(\alpha) = 0$ and $\alpha \equiv \alpha_1 \pmod{p}$. Thus we just need to show such a sequence exists. By assumption, α_1 exists. Now we find α_2. We have

$$\alpha_2 = \alpha_1 + d_1 p$$

for some d_1. Then

$$f(\alpha_2) = f(\alpha_1) + f'(\alpha_1) d_1 p + O(p^2),$$

where $O(p^2)$ denotes some element of $p^2\mathbb{Z}_p$. Thus

$$d_1 p f'(\alpha_1) + f(\alpha_1) \equiv 0 \pmod{p^2}.$$

Note that $f(\alpha_1) \equiv 0 \pmod{p}$, so $f(\alpha_1) = p\beta$ for some $\beta \in \mathbb{Z}_p$. Dividing by p, we get that

$$d_1 f'(\alpha_1) + \beta \equiv 0 \pmod{p},$$

so

$$d_1 \equiv -\beta f'(\alpha_1)^{-1} \pmod{p},$$

which is defined because $f'(\alpha_1) \neq 0 \pmod{p}$. Similar calculations gives a_3 from a_2, a_4 from a_3, and so on. □

3 Main Application

Now we are ready to look at our main application. In this application, we look at the function that rotates a complex number by a specific angle.

Definition 20. A planar rotation is a map $\mathbb{R}^2 \to \mathbb{R}^2$ given by

$$\begin{bmatrix} x \\ y \end{bmatrix} \quad \mapsto \quad \begin{bmatrix} \cos(2\pi v) & -\sin(2\pi v) \\ \sin(2\pi v) & \cos(2\pi v) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix},$$

which describes a rotation by an angle of $2\pi v$ about the origin. We say that v is the *rotation number* of this rotation.

We will investigate the case when v is irrational (if v is rational, then this map is periodic). Let

$$A = \begin{bmatrix} \lambda & -1 \\ 1 & 0 \end{bmatrix} \qquad \lambda = 2\cos(2\pi v)$$

Suppose $v \neq 0, \frac{1}{2}$, so λ is not an integer (these are trivial cases that do not interest us). If

$$C = \begin{bmatrix} 1 & -\cos(2\pi v) \\ 0 & \sin(2\pi v) \end{bmatrix},$$

and J is the matrix corresponding to the planar rotation with rotation number v, we can check that $A = CJC^{-1}$, i.e. J and A are conjugate. This means that the dynamics of A and J have the same orbit structure (so if we understand the dynamics of A we understand the dynamics of J and vice versa). The invariant sets of A are the ellipses

$$x^2 - \lambda xy + y^2 = c.$$

We have the following result.

Lemma 21 (Niven's Theorem)**.** *If λ is rational (and not equal to an integer), then v is irrational.*

Proof. We omit the proof as it is quite complicated. □

Now, we perturb the linear mapping defined by A by discretizing the space. Consider the lattice map

$$\Psi \colon \mathbb{Z}^2 \to \mathbb{Z}^2 \qquad (x, y) \mapsto (\lfloor \lambda x \rfloor - y, x) \qquad \lambda = 2\cos(2\pi v).$$

Then Ψ is just the action of the matrix A but with rounding. Note that the mapping

$$(x, y) \mapsto (y, \lfloor \lambda y \rfloor - x)$$

is an inverse of Ψ, so Ψ is invertible. If there are a finite number of points in the orbit of Ψ, then the orbit must be periodic. If there are an infinite number of points, the orbit must spiral off to infinity because there are finitely many lattice points within any given radius from the origin. Thus the orbits of Ψ are either periodic or spiral off to infinity.

The following image shows a portion of the periodic orbit of Ψ when $\lambda = \frac{1}{2}$. Here is a link to a Google Sheets that generates a graph of the periodic orbit of Ψ for different values of λ and starting points.

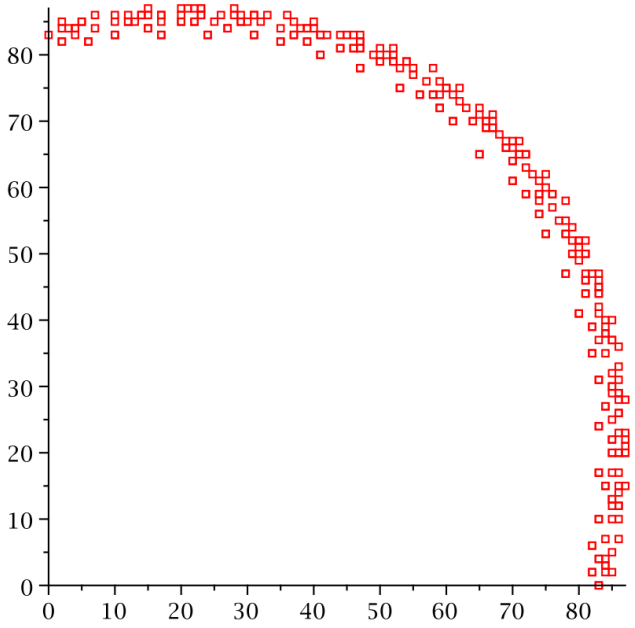


Figure 1: A portion of a periodic orbit of the map Ψ, for $\lambda = \frac{1}{2}$. The period is 696.

We will consider the family of parameter values

$$\lambda = \frac{q}{p^n} \qquad n \geqslant 1, \quad |q| < 2p^n, \quad \gcd(q,p) = 1.$$

Since λ is rational, it follows that the rotation number v is irrational. Consider the polynomial $f(x) = x^2 - qx + p^{2n}$. Then

$$f(x) \equiv x(x-q) \pmod{p^{2n}}, \qquad f'(x) = 2x - q,$$

and we have

$$\begin{aligned} f(0) = f(q) &\equiv 0 \pmod{p^{2n}}, \\ f'(0) &\equiv -q \pmod{p^{2n}}, \\ f'(q) &\equiv q \pmod{p^{2n}}. \end{aligned}$$

Since p and q are relatively prime, it follows from Hensel's lemma that f has two distinct roots θ and $\overline{\theta}$ in $\mathbb{Z}_p$ such that

$$\theta \equiv 0 \pmod{p^{2n}} \qquad \text{and} \qquad \overline{\theta} \equiv q \pmod{p^{2n}}.$$

Lemma 22. *We have* $|\theta|_p = \frac{1}{p^{2n}}$ *and* $|\overline{\theta}|_p = 1$.

Proof. Since $\theta \equiv 0 \pmod{p^{2n}}$, we have $p^{2n} \mid \theta$, which implies $|\theta|_p \leqslant \frac{1}{p^{2n}}$. Suppose, for the sake of contradiction, that $|\theta|_p < \frac{1}{p^{2n}}$. This means that $p^{2n+1} \mid \theta$, and it follows that

$$p^{2n+1} \mid \theta^2 + q\theta = f(\theta) - p^{2n} = -p^{2n},$$

which is obviously false. Thus $|\theta|_p = \frac{1}{p^{2n}}$. Then $\overline{\theta} \equiv q \pmod{p^{2n}}$ implies $\overline{\theta} \equiv q \pmod{p}$. Since $\gcd(p,q) = 1$, it follows that $p \nmid q$, so $|q|_p = 1$. It follows that $|\overline{\theta}|_p = 1$. □

Then it follows that $\frac{\theta}{p^n} \in p^n\mathbb{Z}_p$. Thus we can define the map

$$\mathfrak{L}\colon \mathbb{Z}^2 \to \mathbb{Z}_p, \qquad (x, y) \mapsto x - y \cdot \frac{\theta}{p^n}.$$

Then $\mathfrak{L}$ is injective. Let $\mathcal{L}$ be the image of $\mathbb{Z}^2$ under $\mathfrak{L}$, i.e. $\mathcal{L} = \mathfrak{L}(\mathbb{Z}^2) \subset \mathbb{Z}_p$. Then $\mathcal{L}$ is an additive subgroup of $\mathbb{Z}_p$. Define the map

$$\Psi^*\colon \mathcal{L} \to \mathcal{L}, \qquad \Psi^* = \mathfrak{L} \circ \Psi \circ \mathfrak{L}^{-1}.$$

Then Ψ^* is conjugate to Ψ, which essentially means that if we understand the dynamics of Ψ^* we understand the dynamics of Ψ and vice versa. To characterize the map Ψ^*, we first define the p-adic shift mapping.

Definition 23. The *p-adic shift*

$$\sigma : \mathbb{Z}_p \to \mathbb{Z}_p$$

is defined as follows. Given a p-adic integer

$$z = b_0 + b_1 p + b_2 p^2 + \cdots,$$

where the $b_i \in \{0, 1, \ldots, p-1\}$, we let

$$\sigma(z) = b_1 + b_2 p + b_3 p^2 + \cdots.$$

We immediately see that if σ^k denotes the k-fold iterate of σ, then we have that

$$\sigma^k(z) = b_k + b_{k+1} p + b_{k+2} p^2 + \cdots.$$

Moreover, for $x \in \mathbb{Z}$, it is the case that

$$\sigma^k(x) = \left\lfloor \frac{x}{p^k} \right\rfloor.$$

One of the most basic properties of the shift is that it is continuous as a function of $\mathbb{Z}_p$. Indeed, if $|x - y|_p < 1/p^{k+1}$, it is not hard to see that $|S(x) - S(y)|_p < 1/p^k$.
Finally, given $x \in \mathbb{Z}$, we define the integer $c(x)$ by

$$\frac{qx - c(x)}{p^n} = \left\lfloor \frac{qx}{p^n} \right\rfloor.$$

Definition 24. Let $\Gamma : S \to \mathbb{Z}_p$ be a map from a subset S of $\mathbb{Z}_p$ to $\mathbb{Z}_p$. We say Γ can be *extended continuously* to $\mathbb{Z}_p$ if we can expand the domain of Γ to take in all values of $\mathbb{Z}_p$ such that Γ is continuous, i.e. for any $z_0 \in \mathbb{Z}_p$ and $\Delta \in \mathbb{Q}_{>0}$, there exists $\epsilon \in \mathbb{Q}_{>0}$ such that for all $z \in \mathbb{Z}_p$ with $|z - z_0|_p \leqslant \epsilon$, we have $|\Gamma(z) - \Gamma(z_0)|_p \leqslant \Delta$.

Now we present our main result:

Theorem 25. *We can take the map Ψ^* and extend it continuously so that it is defined for all values in $\mathbb{Z}_p$, and is given by*

$$\Psi^*\colon \mathbb{Z}_p \to \mathbb{Z}_p \qquad z \mapsto \sigma^n(\overline{\theta} z).$$

Proof. Because θ and $\overline{\theta}$ are the roots of $f(z)$, we have that

$$\theta + \overline{\theta} = q \quad \text{and} \quad \theta\overline{\theta} = p^{2n}.$$

Let $z = \mathfrak{L}(x, y) = x - y\frac{\theta}{p^n}$. Then

$$\begin{aligned}\Phi^*(z) &= \mathfrak{L}\left(\left\lfloor \frac{qx}{p^n} \right\rfloor - y, x\right) \\ &= \left\lfloor \frac{qx}{p^n} \right\rfloor - y - x\frac{\theta}{p^n} \\ &= \frac{1}{p^n}(x(q-\theta) - p^n y - c(x)) \\ &= \frac{1}{p^n}(x\overline{\theta} - \frac{\theta\overline{\theta}}{p^n}y - c(x)) \\ &= \frac{1}{p^n}(\overline{\theta}z - c(x)).\end{aligned}$$

Note that $y\frac{\theta}{p^n} \in p^n\mathbb{Z}_p$ since $\frac{\theta}{p^n}$. Thus

$$qx \equiv q_z \equiv \overline{\theta}z \pmod{p^n}.$$

Thus

$$\Psi^*(z) = \sigma^n(\overline{\theta}z).$$

If $z^{(k)} \to z$ is a Cauchy sequence in $\mathcal{L}$, then so is $\sigma^n(\overline{\theta}z^{(k)})$, and since $\mathcal{L}$ is dense in $\mathbb{Z}_p$, we can extend Ψ^* to the whole of $\mathbb{Z}_p$. □

This shows that in essence, the dynamics induced by the round-off errors can be explained quite nicely using p-adic numbers and the p-adic shift!

References

[1] Vivaldi, Franco. 2011. *An Introduction to Arithmetic Dynamics.* Summer School on Dynamical Systems, Mathematisches Institut, Georg-August Universität Göttingen.

[2] Houston-Edwards, Kelsey. *Mathematicians Set Numbers in Motion to Unlock Their Secrets.* Quanta Magazine, 22 Feb. 2021, available online at Quanta Magazine.

[3] Silverman, Joseph H. 2010. *Lecture Notes on Arithmetic Dynamics.* Arizona Winter School, March 13–17, 2010. pdf

[4] Silverman, Joseph H. 2022. *Arithmetic Dynamics: A Survey.* International Congress of Mathematicians, Tuesday July 12, 2022 Slides

[5] Rubinstein-Salzedo, Simon. 2023. *Number Theory.* Lecture notes for Euler Circle, Fall 2023.