Examples in dynamical systems continued fraction

Pierre Arnoux

Lecture 3

The Pigeonhole Principle

- **Theorem (Pigeonhole / Dirichlet)**: If $n+1$ objects are distributed among n boxes, some box contains at least 2 objects.

The Pigeonhole Principle

- **Theorem (Pigeonhole / Dirichlet)**: If $n+1$ objects are distributed among n boxes, some box contains at least 2 objects.
- **Proof**: If every box had at most 1 object, the total would be $\leq n$. Contradiction.

The Pigeonhole Principle

- **Theorem (Pigeonhole / Dirichlet)**: If $n+1$ objects are distributed among n boxes, some box contains at least 2 objects.
- **Proof**: If every box had at most 1 object, the total would be $\leq n$. Contradiction.
- **Generalized form**: If $mn+1$ objects are distributed among n boxes, some box contains at least $m+1$ objects.

The Pigeonhole Principle

- **Theorem (Pigeonhole / Dirichlet)**: If $n+1$ objects are distributed among n boxes, some box contains at least 2 objects.
- **Proof**: If every box had at most 1 object, the total would be $\leq n$. Contradiction.
- **Generalized form**: If $mn+1$ objects are distributed among n boxes, some box contains at least $m+1$ objects.
- Seemingly trivial , yet it is one of the most powerful tools in mathematics

The Pigeonhole Principle

- **Theorem (Pigeonhole / Dirichlet)**: If $n+1$ objects are distributed among n boxes, some box contains at least 2 objects.
- **Proof**: If every box had at most 1 object, the total would be $\leq n$. Contradiction.
- **Generalized form**: If $mn+1$ objects are distributed among n boxes, some box contains at least $m+1$ objects.
- Seemingly trivial , yet it is one of the most powerful tools in mathematics
- The art lies in choosing the right objects and boxes

The Pigeonhole Principle

- **Theorem (Pigeonhole / Dirichlet)**: If $n+1$ objects are distributed among n boxes, some box contains at least 2 objects.
- **Proof**: If every box had at most 1 object, the total would be $\leq n$. Contradiction.
- **Generalized form**: If $mn+1$ objects are distributed among n boxes, some box contains at least $m+1$ objects.
- Seemingly trivial , yet it is one of the most powerful tools in mathematics
- The art lies in choosing the right objects and boxes
- Also known as the *Dirichlet box principle* or *Schubfachprinzip*

Dirichlet's approximation theorem

- **Theorem (Dirichlet, 1842)**: For any real α and integer $N \geq 1$, there exists $1 \leq q \leq N$ and $p \in \mathbb{Z}$ such that $\left|\alpha - \frac{p}{q}\right| < \frac{1}{qN}$.

Dirichlet's approximation theorem

- **Theorem (Dirichlet, 1842)**: For any real α and integer $N \geq 1$, there exists $1 \leq q \leq N$ and $p \in \mathbb{Z}$ such that $\left|\alpha - \frac{p}{q}\right| < \frac{1}{qN}$.
- **Proof**: Consider the $N+1$ fractional parts $\{0 \cdot \alpha\}, \{\alpha\}, \{2\alpha\}, \ldots, \{N\alpha\} \in [0, 1)$.

Dirichlet's approximation theorem

- **Theorem (Dirichlet, 1842)**: For any real α and integer $N \geq 1$, there exists $1 \leq q \leq N$ and $p \in \mathbb{Z}$ such that $\left|\alpha - \frac{p}{q}\right| < \frac{1}{qN}$.
- **Proof**: Consider the $N+1$ fractional parts $\{0 \cdot \alpha\}, \{\alpha\}, \{2\alpha\}, \ldots, \{N\alpha\} \in [0,1)$.
- Divide $[0,1)$ into N equal subintervals $\left[\frac{i}{N}, \frac{i+1}{N}\right)$, $i = 0, \ldots, N-1$.

Dirichlet's approximation theorem

- **Theorem (Dirichlet, 1842)**: For any real α and integer $N \geq 1$, there exists $1 \leq q \leq N$ and $p \in \mathbb{Z}$ such that $\left|\alpha - \frac{p}{q}\right| < \frac{1}{qN}$.
- **Proof**: Consider the $N+1$ fractional parts $\{0 \cdot \alpha\}, \{\alpha\}, \{2\alpha\}, \ldots, \{N\alpha\} \in [0,1)$.
- Divide $[0,1)$ into N equal subintervals $\left[\frac{i}{N}, \frac{i+1}{N}\right)$, $i = 0, \ldots, N-1$.
- By pigeonhole: some two fractional parts $\{j\alpha\}$ and $\{k\alpha\}$, $j > k$, fall in the same subinterval.

Dirichlet's approximation theorem

- **Theorem (Dirichlet, 1842)**: For any real α and integer $N \geq 1$, there exists $1 \leq q \leq N$ and $p \in \mathbb{Z}$ such that $\left|\alpha - \frac{p}{q}\right| < \frac{1}{qN}$.
- **Proof**: Consider the $N+1$ fractional parts $\{0 \cdot \alpha\}, \{\alpha\}, \{2\alpha\}, \ldots, \{N\alpha\} \in [0,1)$.
- Divide $[0,1)$ into N equal subintervals $\left[\frac{i}{N}, \frac{i+1}{N}\right)$, $i = 0, \ldots, N-1$.
- By pigeonhole: some two fractional parts $\{j\alpha\}$ and $\{k\alpha\}$, $j > k$, fall in the same subinterval.
- Hence $|\{j\alpha\} - \{k\alpha\}| < 1/N$, i.e. $|(j-k)\alpha - p| < 1/N$ for some $p \in \mathbb{Z}$.

Dirichlet's approximation theorem

- **Theorem (Dirichlet, 1842)**: For any real α and integer $N \geq 1$, there exists $1 \leq q \leq N$ and $p \in \mathbb{Z}$ such that $\left|\alpha - \frac{p}{q}\right| < \frac{1}{qN}$.
- **Proof**: Consider the $N+1$ fractional parts $\{0 \cdot \alpha\}, \{\alpha\}, \{2\alpha\}, \ldots, \{N\alpha\} \in [0,1)$.
- Divide $[0,1)$ into N equal subintervals $\left[\frac{i}{N}, \frac{i+1}{N}\right)$, $i = 0, \ldots, N-1$.
- By pigeonhole: some two fractional parts $\{j\alpha\}$ and $\{k\alpha\}$, $j > k$, fall in the same subinterval.
- Hence $|\{j\alpha\} - \{k\alpha\}| < 1/N$, i.e. $|(j-k)\alpha - p| < 1/N$ for some $p \in \mathbb{Z}$.
- Set $q = j - k \leq N$: we get $|\alpha - p/q| < 1/(qN)$. $\square$

Dirichlet's theorem: consequences

- **Corollary**: Every irrational α has infinitely many rational approximations p/q with $|\alpha - p/q| < 1/q^2$.

Dirichlet's theorem: consequences

- **Corollary**: Every irrational α has infinitely many rational approximations p/q with $|\alpha - p/q| < 1/q^2$.
- **Proof**: Apply Dirichlet for $N = 1, 2, 3, \ldots$ to get infinitely many (p_N, q_N) with $q_N \leq N$.

Dirichlet's theorem: consequences

- **Corollary**: Every irrational α has infinitely many rational approximations p/q with $|\alpha - p/q| < 1/q^2$.
- **Proof**: Apply Dirichlet for $N = 1, 2, 3, \ldots$ to get infinitely many (p_N, q_N) with $q_N \leq N$.
- Since α is irrational, $\alpha \neq p/q$ for all p/q, hence infinitely many are distinct.

Dirichlet's theorem: consequences

- **Corollary**: Every irrational α has infinitely many rational approximations p/q with $|\alpha - p/q| < 1/q^2$.
- **Proof**: Apply Dirichlet for $N = 1, 2, 3, \ldots$ to get infinitely many (p_N, q_N) with $q_N \leq N$.
- Since α is irrational, $\alpha \neq p/q$ for all p/q, hence infinitely many are distinct.
- **A remarkable fact**: This bound $1/q^2$ is sharp for many numbers: Roth's theorem 1955 shows $1/q^{2+\varepsilon}$ fails for all algebraic irrationals (but there are Liouville numbers which are very well approximated)

Dirichlet's theorem: consequences

- **Corollary**: Every irrational α has infinitely many rational approximations p/q with $|\alpha - p/q| < 1/q^2$.
- **Proof**: Apply Dirichlet for $N = 1, 2, 3, \ldots$ to get infinitely many (p_N, q_N) with $q_N \leq N$.
- Since α is irrational, $\alpha \neq p/q$ for all p/q, hence infinitely many are distinct.
- **A remarkable fact**: This bound $1/q^2$ is sharp for many numbers: Roth's theorem 1955 shows $1/q^{2+\varepsilon}$ fails for all algebraic irrationals (but there are Liouville numbers which are very well approximated)
- all *convergents* of the continued fraction of α give Best approximations (see later)

Dirichlet's theorem: consequences

- **Corollary**: Every irrational α has infinitely many rational approximations p/q with $|\alpha - p/q| < 1/q^2$.
- **Proof**: Apply Dirichlet for $N = 1, 2, 3, \ldots$ to get infinitely many (p_N, q_N) with $q_N \leq N$.
- Since α is irrational, $\alpha \neq p/q$ for all p/q, hence infinitely many are distinct.
- **A remarkable fact**: This bound $1/q^2$ is sharp for many numbers: Roth's theorem 1955 shows $1/q^{2+\varepsilon}$ fails for all algebraic irrationals (but there are Liouville numbers which are very well approximated)
- all *convergents* of the continued fraction of α give Best approximations (see later)
- The pigeonhole principle, applied to fractional parts, gives us a bridge to number theory and dynamics

A remark

- Trying to minimize $|\alpha - \frac{p}{q}|$ and $|q\alpha - p|$ are related problems:

A remark

- Trying to minimize $|\alpha - \frac{p}{q}|$ and $|q\alpha - p|$ are related problems:
- $|q\alpha - p| < \epsilon$ is equivalent to $|\alpha - \frac{p}{q}| < \frac{\epsilon}{q}$.

A remark

- Trying to minimize $|\alpha - \frac{p}{q}|$ and $|q\alpha - p|$ are related problems:
- $|q\alpha - p| < \epsilon$ is equivalent to $|\alpha - \frac{p}{q}| < \frac{\epsilon}{q}$.
- The second inequality gives us good rational approximations of a fixed irrational α.

A remark

- Trying to minimize $|\alpha - \frac{p}{q}|$ and $|q\alpha - p|$ are related problems:
- $|q\alpha - p| < \epsilon$ is equivalent to $|\alpha - \frac{p}{q}| < \frac{\epsilon}{q}$.
- The second inequality gives us good rational approximations of a fixed irrational α.
- The first inequality tells us the best return of the orbit of 0 by the rotation of angle α to the origin.

The usual continued fraction: the goal

- **Problem**: write an irrational $x \in (0,1)$ as an infinite continued fraction

$$x = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cdots}} = [0; a_1, a_2, a_3, \ldots], \qquad a_n \in \mathbb{N}^*.$$

The usual continued fraction: the goal

- **Problem**: write an irrational $x \in (0,1)$ as an infinite continued fraction

$$x = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cdots}} = [0; a_1, a_2, a_3, \ldots], \qquad a_n \in \mathbb{N}^*.$$

- **Computation**: subtract the integer part, invert, and repeat:

$$x_0 = x, \qquad a_{n+1} = \left\lfloor \frac{1}{x_n} \right\rfloor, \qquad x_{n+1} = \frac{1}{x_n} - a_{n+1} \in (0,1).$$

The usual continued fraction: the goal

- **Problem**: write an irrational $x \in (0,1)$ as an infinite continued fraction

$$x = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cdots}} = [0; a_1, a_2, a_3, \ldots], \qquad a_n \in \mathbb{N}^*.$$

- **Computation**: subtract the integer part, invert, and repeat:

$$x_0 = x, \qquad a_{n+1} = \left\lfloor \frac{1}{x_n} \right\rfloor, \qquad x_{n+1} = \frac{1}{x_n} - a_{n+1} \in (0,1).$$

- Why it works: $\frac{1}{x_n} = a_{n+1} + x_{n+1}$, so after n steps $x = [0; a_1, \ldots, a_n + x_n]$.

The usual continued fraction: the goal

- **Problem**: write an irrational $x \in (0,1)$ as an infinite continued fraction

$$x = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cdots}} = [0; a_1, a_2, a_3, \ldots], \qquad a_n \in \mathbb{N}^*.$$

- **Computation**: subtract the integer part, invert, and repeat:

$$x_0 = x, \qquad a_{n+1} = \left\lfloor \frac{1}{x_n} \right\rfloor, \qquad x_{n+1} = \frac{1}{x_n} - a_{n+1} \in (0,1).$$

- Why it works: $\frac{1}{x_n} = a_{n+1} + x_{n+1}$, so after n steps $x = [0; a_1, \ldots, a_n + x_n]$.
- If x is irrational, every x_n is irrational and nonzero: the algorithm never stops. If x is rational it stops (some $x_n = 0$): this is Euclid's algorithm on numerator and denominator.

The usual continued fraction: examples

- $x = \sqrt{2} - 1$: $\frac{1}{x} = \sqrt{2} + 1 = 2 + x$. So $a_1 = 2$ and $x_1 = x$: the algorithm repeats forever, $x = [0; 2, 2, 2, \ldots]$.

The usual continued fraction: examples

- $x = \sqrt{2} - 1$: $\frac{1}{x} = \sqrt{2} + 1 = 2 + x$. So $a_1 = 2$ and $x_1 = x$: the algorithm repeats forever, $x = [0; 2, 2, 2, \ldots]$.
- $x = \frac{\sqrt{5}-1}{2}$: $\frac{1}{x} = \frac{\sqrt{5}+1}{2} = 1 + x$, so $x = [0; 1, 1, 1, \ldots]$: the golden ratio, the hardest number to approximate.

The usual continued fraction: examples

- $x = \sqrt{2} - 1$: $\frac{1}{x} = \sqrt{2} + 1 = 2 + x$. So $a_1 = 2$ and $x_1 = x$: the algorithm repeats forever, $x = [0; 2, 2, 2, \ldots]$.
- $x = \frac{\sqrt{5}-1}{2}$: $\frac{1}{x} = \frac{\sqrt{5}+1}{2} = 1 + x$, so $x = [0; 1, 1, 1, \ldots]$: the golden ratio, the hardest number to approximate.
- $x = \pi - 3 = 0.14159265\ldots$: $\frac{1}{x} = 7.0625\ldots$, then $15.9966\ldots$, then $1.0034\ldots$, then $292.63\ldots$

$$\pi - 3 = [0; 7, 15, 1, 292, 1, 1, 1, 2, 1, 3, \ldots]$$

The usual continued fraction: examples

- $x = \sqrt{2} - 1$: $\frac{1}{x} = \sqrt{2} + 1 = 2 + x$. So $a_1 = 2$ and $x_1 = x$: the algorithm repeats forever, $x = [0; 2, 2, 2, \ldots]$.
- $x = \frac{\sqrt{5}-1}{2}$: $\frac{1}{x} = \frac{\sqrt{5}+1}{2} = 1 + x$, so $x = [0; 1, 1, 1, \ldots]$: the golden ratio, the hardest number to approximate.
- $x = \pi - 3 = 0.14159265\ldots$: $\frac{1}{x} = 7.0625\ldots$, then $15.9966\ldots$, then $1.0034\ldots$, then $292.63\ldots$

$$\pi - 3 = [0; 7, 15, 1, 292, 1, 1, 1, 2, 1, 3, \ldots]$$

- $e - 2 = [0; 1, 2, 1, 1, 4, 1, 1, 6, 1, 1, 8, \ldots]$ has a pattern; no pattern is known for π.

The usual continued fraction: vocabulary

- **Remainders**: $x_n = T^n(x) \in (0,1)$, where T is the Gauss map (with $x_0 = x$).

The usual continued fraction: vocabulary

- **Remainders**: $x_n = T^n(x) \in (0,1)$, where T is the Gauss map (with $x_0 = x$).
- **Partial quotients**: $a_n = \lfloor 1/x_{n-1} \rfloor \in \mathbb{N}^*$, so that $\frac{1}{x_{n-1}} = a_n + x_n$.

The usual continued fraction: vocabulary

- **Remainders**: $x_n = T^n(x) \in (0,1)$, where T is the Gauss map (with $x_0 = x$).
- **Partial quotients**: $a_n = \lfloor 1/x_{n-1} \rfloor \in \mathbb{N}^*$, so that $\frac{1}{x_{n-1}} = a_n + x_n$.
- **Convergents**: the rationals $\frac{p_n}{q_n} = [0; a_1, \ldots, a_n]$ obtained by truncating the expansion.

The usual continued fraction: vocabulary

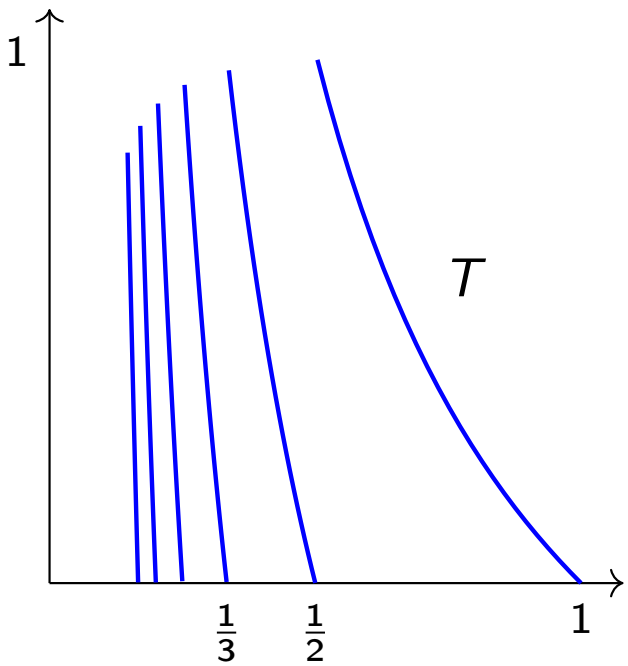
- **Remainders**: $x_n = T^n(x) \in (0,1)$, where T is the Gauss map (with $x_0 = x$).
- **Partial quotients**: $a_n = \lfloor 1/x_{n-1} \rfloor \in \mathbb{N}^*$, so that $\frac{1}{x_{n-1}} = a_n + x_n$.
- **Convergents**: the rationals $\frac{p_n}{q_n} = [0; a_1, \ldots, a_n]$ obtained by truncating the expansion.
- The remainder x_n carries all the information on the tail: $x = [0; a_1, \ldots, a_n + x_n]$.

The usual continued fraction: vocabulary

- **Remainders**: $x_n = T^n(x) \in (0,1)$, where T is the Gauss map (with $x_0 = x$).
- **Partial quotients**: $a_n = \lfloor 1/x_{n-1} \rfloor \in \mathbb{N}^*$, so that $\frac{1}{x_{n-1}} = a_n + x_n$.
- **Convergents**: the rationals $\frac{p_n}{q_n} = [0; a_1, \ldots, a_n]$ obtained by truncating the expansion.
- The remainder x_n carries all the information on the tail: $x = [0; a_1, \ldots, a_n + x_n]$.
- Example, $x = \pi$:

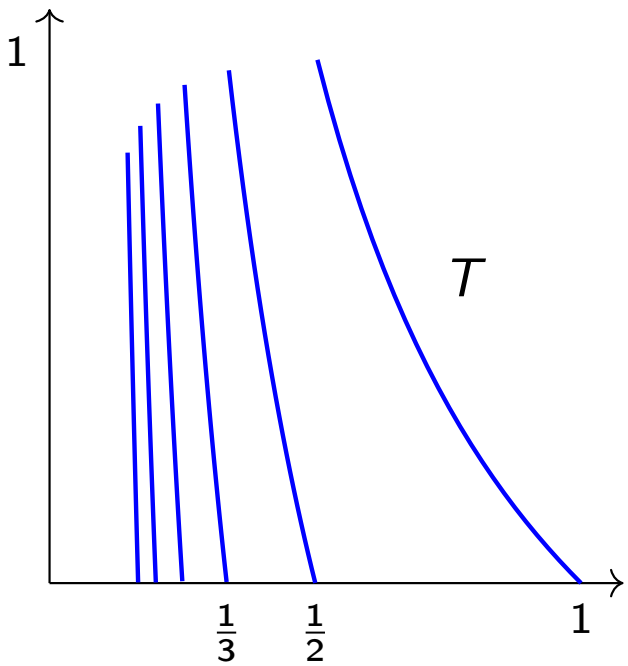
n	0	1	2	3	4
a_n	3	7	15	1	292
p_n/q_n	$\frac{3}{1}$	$\frac{22}{7}$	$\frac{333}{106}$	$\frac{355}{113}$	$\frac{103993}{33102}$

The usual continued fraction: the Gauss map



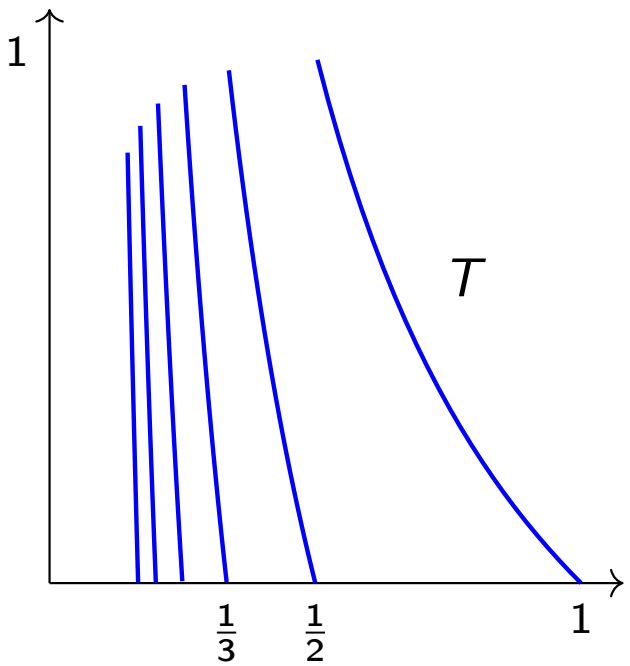
- **Gauss map**:
  $T(x) = \{1/x\} = \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor$ on $(0, 1)$.

The usual continued fraction: the Gauss map



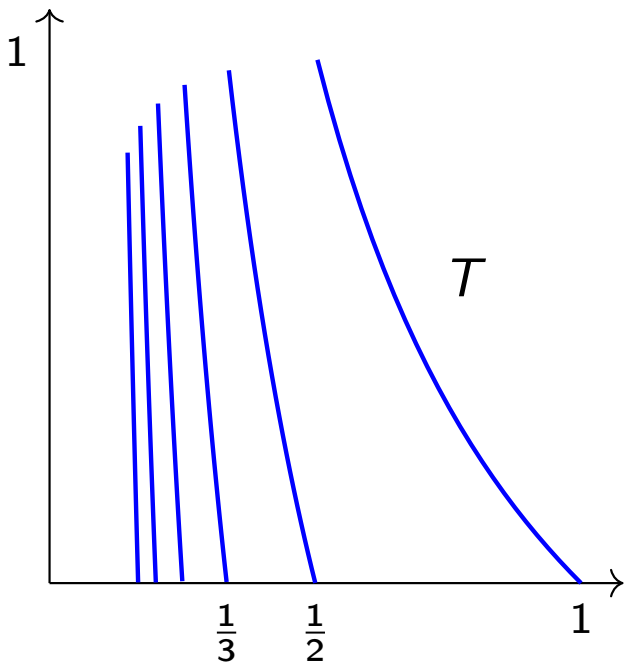
- **Gauss map**:
  $T(x) = \{1/x\} = \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor$ on $(0, 1)$.
- Infinitely many branches: on $I_k = \left(\frac{1}{k+1}, \frac{1}{k}\right]$, $T(x) = \frac{1}{x} - k$ is decreasing, onto $[0, 1)$.

The usual continued fraction: the Gauss map



- **Gauss map**: $T(x) = \{1/x\} = \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor$ on $(0,1)$.
- Infinitely many branches: on $I_k = \left(\frac{1}{k+1}, \frac{1}{k}\right]$, $T(x) = \frac{1}{x} - k$ is decreasing, onto $[0,1)$.
- **Coding map**: $a(x) = \lfloor 1/x \rfloor = k$ on I_k. The partial quotients are $a_n = a(T^{n-1}x)$.

The usual continued fraction: the Gauss map



- **Gauss map**: $T(x) = \{1/x\} = \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor$ on $(0,1)$.
- Infinitely many branches: on $I_k = \left(\frac{1}{k+1}, \frac{1}{k}\right]$, $T(x) = \frac{1}{x} - k$ is decreasing, onto $[0,1)$.
- **Coding map**: $a(x) = \lfloor 1/x \rfloor = k$ on I_k. The partial quotients are $a_n = a(T^{n-1}x)$.
- The digits are the **coding of the orbit** of x by the partition (I_k): we have the coding of a dynamical system!

The usual continued fraction: formulas (1)

- **Recurrence**: with $p_{-1} = 1,\ q_{-1} = 0,\ p_0 = 0,\ q_0 = 1,$

$$p_n = a_n p_{n-1} + p_{n-2}, \qquad q_n = a_n q_{n-1} + q_{n-2}.$$

The usual continued fraction: formulas (1)

- **Recurrence**: with $p_{-1} = 1,\ q_{-1} = 0,\ p_0 = 0,\ q_0 = 1,$

$$p_n = a_n p_{n-1} + p_{n-2}, \qquad q_n = a_n q_{n-1} + q_{n-2}.$$

- **Matrix form**:

$$M_n = \begin{pmatrix} p_{n-1} & p_n \\ q_{n-1} & q_n \end{pmatrix} = M_{n-1} \begin{pmatrix} 0 & 1 \\ 1 & a_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & a_1 \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & a_n \end{pmatrix}$$

The usual continued fraction: formulas (1)

- **Recurrence**: with $p_{-1} = 1,\ q_{-1} = 0,\ p_0 = 0,\ q_0 = 1,$

$$p_n = a_n p_{n-1} + p_{n-2}, \qquad q_n = a_n q_{n-1} + q_{n-2}.$$

- **Matrix form**:

$$M_n = \begin{pmatrix} p_{n-1} & p_n \\ q_{n-1} & q_n \end{pmatrix} = M_{n-1} \begin{pmatrix} 0 & 1 \\ 1 & a_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & a_1 \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & a_n \end{pmatrix}$$

- **Determinant**: each factor has determinant -1, hence

$$p_{n-1} q_n - p_n q_{n-1} = (-1)^n,$$

and p_n/q_n is in lowest terms.

The usual continued fraction: formulas (1)

- **Recurrence**: with $p_{-1} = 1,\ q_{-1} = 0,\ p_0 = 0,\ q_0 = 1,$

$$p_n = a_n p_{n-1} + p_{n-2}, \qquad q_n = a_n q_{n-1} + q_{n-2}.$$

- **Matrix form**:

$$M_n = \begin{pmatrix} p_{n-1} & p_n \\ q_{n-1} & q_n \end{pmatrix} = M_{n-1} \begin{pmatrix} 0 & 1 \\ 1 & a_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & a_1 \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & a_n \end{pmatrix}$$

- **Determinant**: each factor has determinant -1, hence

$$p_{n-1} q_n - p_n q_{n-1} = (-1)^n,$$

and p_n/q_n is in lowest terms.

- **Tail**: $x = \dfrac{p_n + p_{n-1} x_n}{q_n + q_{n-1} x_n}$: the Möbius map of the matrix M_n applied to the remainder x_n.

The usual continued fraction: formulas (2)

- **Error**: $x - \frac{p_n}{q_n} = \frac{(-1)^n x_n}{q_n(q_n + q_{n-1}x_n)} = \frac{(-1)^n}{q_n(q_{n+1} + q_n x_{n+1})}$.

The usual continued fraction: formulas (2)

- **Error**: $x - \dfrac{p_n}{q_n} = \dfrac{(-1)^n x_n}{q_n(q_n + q_{n-1}x_n)} = \dfrac{(-1)^n}{q_n(q_{n+1} + q_n x_{n+1})}$.
- **Parity and direction**: even convergents are $< x$, odd ones are $> x$: the convergents *alternate* around x, and $\dfrac{p_n}{q_n} - \dfrac{p_{n-1}}{q_{n-1}} = \dfrac{(-1)^{n-1}}{q_n q_{n-1}}$.

The usual continued fraction: formulas (2)

- **Error**: $x - \dfrac{p_n}{q_n} = \dfrac{(-1)^n x_n}{q_n(q_n + q_{n-1}x_n)} = \dfrac{(-1)^n}{q_n(q_{n+1} + q_n x_{n+1})}$.
- **Parity and direction**: even convergents are $< x$, odd ones are $> x$: the convergents *alternate* around x, and $\dfrac{p_n}{q_n} - \dfrac{p_{n-1}}{q_{n-1}} = \dfrac{(-1)^{n-1}}{q_n q_{n-1}}$.
- **Size**: $\dfrac{1}{q_n(q_{n+1} + q_n)} < \left|x - \dfrac{p_n}{q_n}\right| < \dfrac{1}{q_n q_{n+1}} \leq \dfrac{1}{q_n^2}$, hence $|q_n x - p_n| < \dfrac{1}{q_{n+1}}$.

The usual continued fraction: formulas (2)

- **Error**: $x - \dfrac{p_n}{q_n} = \dfrac{(-1)^n x_n}{q_n(q_n + q_{n-1}x_n)} = \dfrac{(-1)^n}{q_n(q_{n+1} + q_n x_{n+1})}$.
- **Parity and direction**: even convergents are $< x$, odd ones are $> x$: the convergents *alternate* around x, and $\dfrac{p_n}{q_n} - \dfrac{p_{n-1}}{q_{n-1}} = \dfrac{(-1)^{n-1}}{q_n q_{n-1}}$.
- **Size**: $\dfrac{1}{q_n(q_{n+1} + q_n)} < \left|x - \dfrac{p_n}{q_n}\right| < \dfrac{1}{q_n q_{n+1}} \leq \dfrac{1}{q_n^2}$, hence $|q_n x - p_n| < \dfrac{1}{q_{n+1}}$.
- So $|x - p_n/q_n| < 1/q_n^2$ for *all* n: this is Dirichlet's corollary again, now with an explicit sequence of approximations.

The usual continued fraction: best approximations

- **Theorem (Lagrange)**: if $n \geq 1$, $1 \leq q < q_{n+1}$ and $p \in \mathbb{Z}$, then $|qx - p| \geq |q_n x - p_n|$.

The usual continued fraction: best approximations

- **Theorem (Lagrange)**: if $n \geq 1$, $1 \leq q < q_{n+1}$ and $p \in \mathbb{Z}$, then $|qx - p| \geq |q_n x - p_n|$.
- The convergents are the *best approximations*: they give the successive records of the closest returns to 0 of the orbit of 0 under the rotation of angle x (see the remark on Dirichlet's theorem).

The usual continued fraction: best approximations

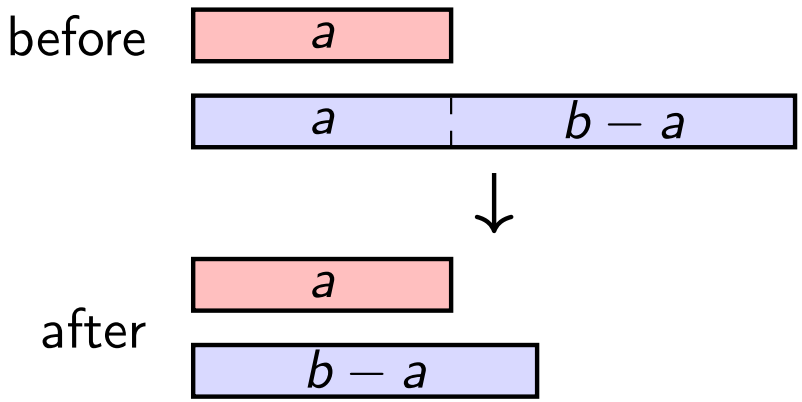
- **Theorem (Lagrange)**: if $n \geq 1$, $1 \leq q < q_{n+1}$ and $p \in \mathbb{Z}$, then $|qx - p| \geq |q_n x - p_n|$.
- The convergents are the *best approximations*: they give the successive records of the closest returns to 0 of the orbit of 0 under the rotation of angle x (see the remark on Dirichlet's theorem).
- **Theorem (Legendre)**: if $\left|x - \frac{p}{q}\right| < \frac{1}{2q^2}$, then $\frac{p}{q}$ is a convergent of x.

The usual continued fraction: best approximations

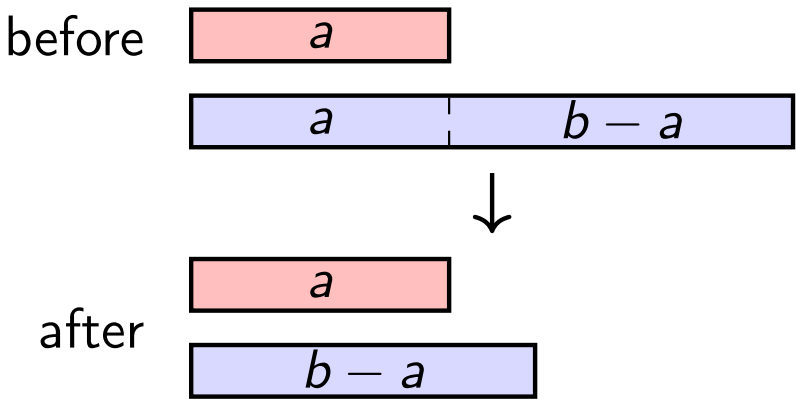
- **Theorem (Lagrange)**: if $n \geq 1$, $1 \leq q < q_{n+1}$ and $p \in \mathbb{Z}$, then $|qx - p| \geq |q_n x - p_n|$.
- The convergents are the *best approximations*: they give the successive records of the closest returns to 0 of the orbit of 0 under the rotation of angle x (see the remark on Dirichlet's theorem).
- **Theorem (Legendre)**: if $\left|x - \frac{p}{q}\right| < \frac{1}{2q^2}$, then $\frac{p}{q}$ is a convergent of x.
- Large partial quotients $\Leftrightarrow$ very good approximations: $a_{n+1} \approx \frac{1}{q_n^2 \, |x - p_n/q_n|}$ (up to a bounded error).

Euclid's algorithm on intervals: the additive version

- Two intervals of lengths $a < b$ (e.g. the two arcs cut by a rotation).



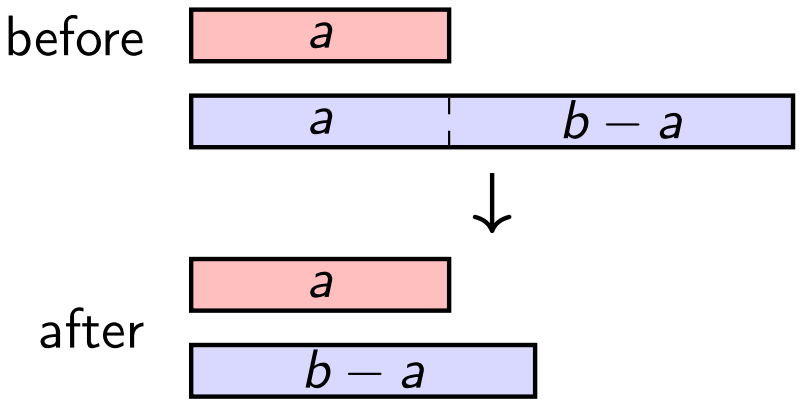
Euclid's algorithm on intervals: the additive version



- Two intervals of lengths $a < b$ (e.g. the two arcs cut by a rotation).
- **Step**: take *one* copy of the shorter interval away from the longer one:

$$(a, b) \longmapsto (a,\ b-a) \quad \text{if } a < b.$$

Euclid's algorithm on intervals: the additive version

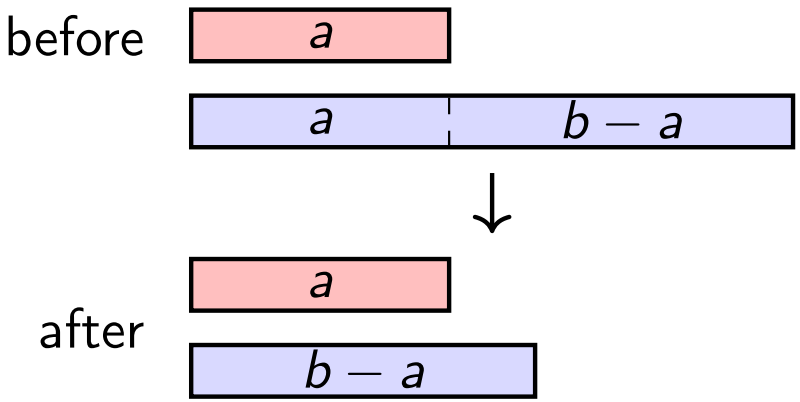


- Two intervals of lengths $a < b$ (e.g. the two arcs cut by a rotation).
- **Step**: take *one* copy of the shorter interval away from the longer one:

$$(a, b) \longmapsto (a,\ b - a) \quad \text{if } a < b.$$

- It stops iff a/b is rational (at some point the two lengths are equal); otherwise it goes on forever.

Euclid's algorithm on intervals: the additive version

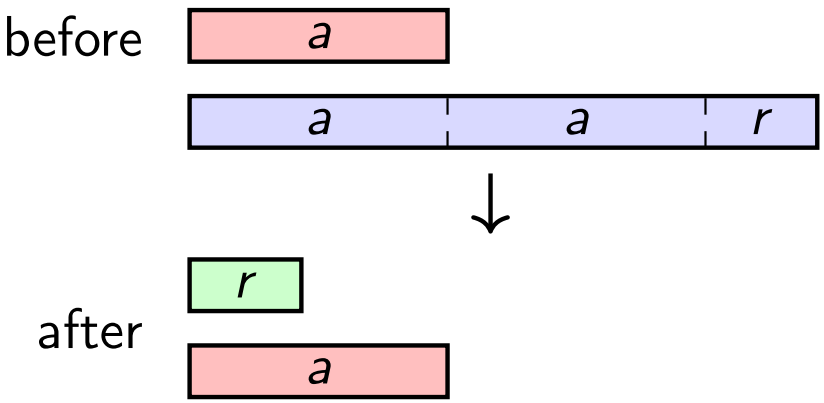


- Two intervals of lengths $a < b$ (e.g. the two arcs cut by a rotation).
- **Step**: take *one* copy of the shorter interval away from the longer one:

$$(a, b) \longmapsto (a,\ b - a) \quad \text{if } a < b.$$

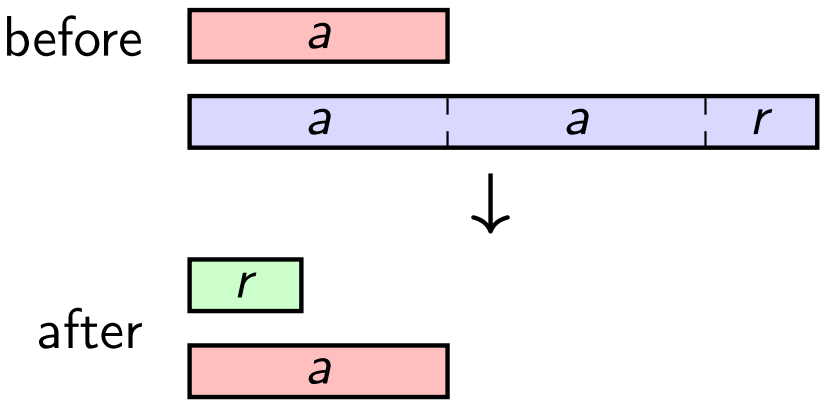
- It stops iff a/b is rational (at some point the two lengths are equal); otherwise it goes on forever.
- The map is linear on each cone $\{a < b\}$, $\{a > b\}$, with matrices $\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$: the inverses of M_1, M_0 (Lecture 2).

Euclid's algorithm on intervals: the multiplicative version



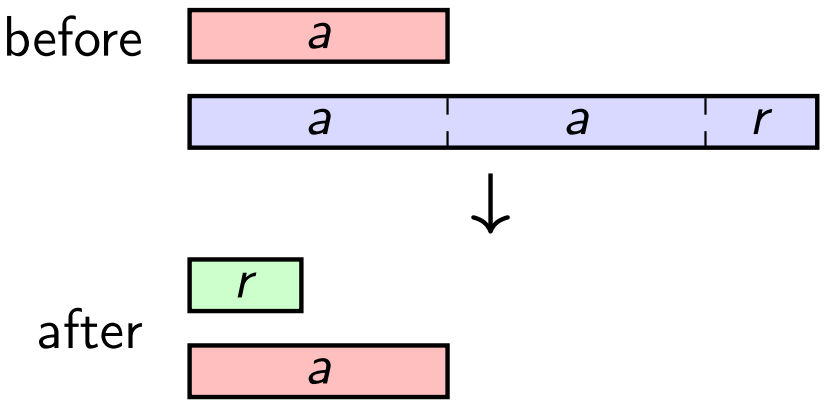
- Take away from the longer interval *as many copies as possible* of the shorter one: $b = ka + r$, with $k = \lfloor b/a \rfloor$ and $0 \leq r < a$.

Euclid's algorithm on intervals: the multiplicative version



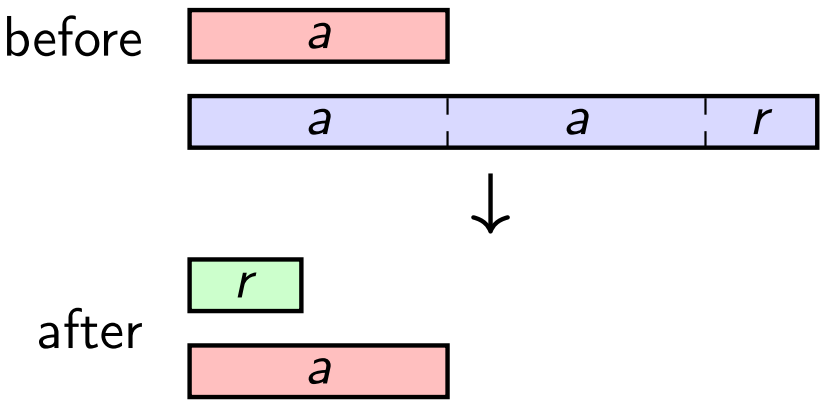
- Take away from the longer interval *as many copies as possible* of the shorter one: $b = ka + r$, with $k = \lfloor b/a \rfloor$ and $0 \leq r < a$.
- $(a, b) \longmapsto (r, a)$: the new pair is again (shorter, longer). The integer k is the **partial quotient**.

Euclid's algorithm on intervals: the multiplicative version



- Take away from the longer interval *as many copies as possible* of the shorter one: $b = ka + r$, with $k = \lfloor b/a \rfloor$ and $0 \leq r < a$.
- $(a, b) \longmapsto (r, a)$: the new pair is again (shorter, longer). The integer k is the **partial quotient**.
- It is k additive steps in a row, $(a, b) \to (a, b - a) \to \cdots \to (a, r)$, followed by a reordering.

Euclid's algorithm on intervals: the multiplicative version



- Take away from the longer interval *as many copies as possible* of the shorter one: $b = ka + r$, with $k = \lfloor b/a \rfloor$ and $0 \leq r < a$.
- $(a, b) \longmapsto (r, a)$: the new pair is again (shorter, longer). The integer k is the **partial quotient**.
- It is k additive steps in a row, $(a, b) \to (a, b - a) \to \cdots \to (a, r)$, followed by a reordering.
- Euclid's algorithm on (a, b) computes the continued fraction of $a/b = [0; k_1, k_2, \ldots]$.

From a piecewise linear map to a projective map

- The additive algorithm is a map $\mathcal{E}$ on pairs of lengths $(\ell_0, \ell_1) \in \mathbb{R}_+^2$, *linear* on each of the cones $\{\ell_0 > \ell_1\}$ and $\{\ell_0 < \ell_1\}$, and **homogeneous**: $\mathcal{E}(\lambda\ell) = \lambda\,\mathcal{E}(\ell)$.

From a piecewise linear map to a projective map

- The additive algorithm is a map $\mathcal{E}$ on pairs of lengths $(\ell_0, \ell_1) \in \mathbb{R}_+^2$, *linear* on each of the cones $\{\ell_0 > \ell_1\}$ and $\{\ell_0 < \ell_1\}$, and **homogeneous**: $\mathcal{E}(\lambda\ell) = \lambda\,\mathcal{E}(\ell)$.
- So only the *shape* of the pair matters: we can always **renormalize**, for instance so that $\ell_0 + \ell_1 = 1$. A pair is then a single number $t = \ell_1 \in (0, 1)$.

From a piecewise linear map to a projective map

- The additive algorithm is a map $\mathcal{E}$ on pairs of lengths $(\ell_0, \ell_1) \in \mathbb{R}_+^2$, *linear* on each of the cones $\{\ell_0 > \ell_1\}$ and $\{\ell_0 < \ell_1\}$, and **homogeneous**: $\mathcal{E}(\lambda\ell) = \lambda\,\mathcal{E}(\ell)$.
- So only the *shape* of the pair matters: we can always **renormalize**, for instance so that $\ell_0 + \ell_1 = 1$. A pair is then a single number $t = \ell_1 \in (0, 1)$.
- Renormalized map: if $t < \frac{1}{2}$, $(1-t, t) \mapsto (1-2t, t)$, of total length $1-t$; if $t > \frac{1}{2}$, $(1-t, t) \mapsto (1-t, 2t-1)$, of total length t. Hence

$$t \longmapsto \frac{t}{1-t} \quad (t < \tfrac{1}{2}), \qquad t \longmapsto \frac{2t-1}{t} \quad (t > \tfrac{1}{2}).$$

From a piecewise linear map to a projective map

- The additive algorithm is a map $\mathcal{E}$ on pairs of lengths $(\ell_0, \ell_1) \in \mathbb{R}_+^2$, *linear* on each of the cones $\{\ell_0 > \ell_1\}$ and $\{\ell_0 < \ell_1\}$, and **homogeneous**: $\mathcal{E}(\lambda\ell) = \lambda\,\mathcal{E}(\ell)$.
- So only the *shape* of the pair matters: we can always **renormalize**, for instance so that $\ell_0 + \ell_1 = 1$. A pair is then a single number $t = \ell_1 \in (0, 1)$.
- Renormalized map: if $t < \frac{1}{2}$, $(1-t, t) \mapsto (1-2t, t)$, of total length $1-t$; if $t > \frac{1}{2}$, $(1-t, t) \mapsto (1-t, 2t-1)$, of total length t. Hence

$$t \longmapsto \frac{t}{1-t} \quad (t < \tfrac{1}{2}), \qquad t \longmapsto \frac{2t-1}{t} \quad (t > \tfrac{1}{2}).$$

- Linear map followed by renormalisation: a **projective map**, i.e. a Möbius map on each branch. It is exactly the recoding map of Lecture 2, t being the frequency of the letter 1.

Additive and multiplicative: Farey and Gauss

- Use now the ratio $x = a/b \in (0, 1)$ of the *shorter* to the *longer* interval, and **reorder** the intervals by length at each step.

Additive and multiplicative: Farey and Gauss

- Use now the ratio $x = a/b \in (0,1)$ of the *shorter* to the *longer* interval, and **reorder** the intervals by length at each step.
- **Additive version**: $(a,b) \mapsto (a, b-a)$, then reorder:

$$F(x) = \begin{cases} \dfrac{x}{1-x} & \text{if } x < \frac{1}{2} \\ \dfrac{1-x}{x} & \text{if } x > \frac{1}{2} \end{cases}$$

the **Farey map**.

Additive and multiplicative: Farey and Gauss

- Use now the ratio $x = a/b \in (0,1)$ of the *shorter* to the *longer* interval, and **reorder** the intervals by length at each step.
- **Additive version**: $(a,b) \mapsto (a, b-a)$, then reorder:

$$F(x) = \begin{cases} \dfrac{x}{1-x} & \text{if } x < \frac{1}{2} \\ \dfrac{1-x}{x} & \text{if } x > \frac{1}{2} \end{cases} \qquad \text{the } \textbf{Farey map}.$$

- **Multiplicative version**: $(a,b) \mapsto (r,a)$ with $b = ka + r$, so

$$x = \frac{a}{b} \longmapsto \frac{r}{a} = \left\{\frac{1}{x}\right\} = T(x), \qquad \text{the } \textbf{Gauss map}.$$

Additive and multiplicative: Farey and Gauss

- Use now the ratio $x = a/b \in (0,1)$ of the *shorter* to the *longer* interval, and **reorder** the intervals by length at each step.
- **Additive version**: $(a,b) \mapsto (a, b-a)$, then reorder:

$$F(x) = \begin{cases} \dfrac{x}{1-x} & \text{if } x < \frac{1}{2} \\ \dfrac{1-x}{x} & \text{if } x > \frac{1}{2} \end{cases} \qquad \text{the } \textbf{Farey map}.$$

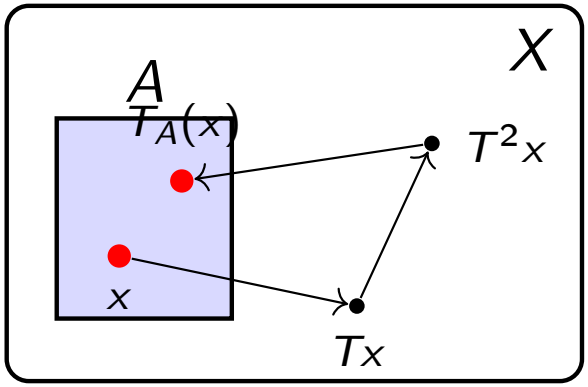
- **Multiplicative version**: $(a,b) \mapsto (r,a)$ with $b = ka + r$, so

$$x = \frac{a}{b} \longmapsto \frac{r}{a} = \left\{\frac{1}{x}\right\} = T(x), \qquad \text{the } \textbf{Gauss map}.$$

- The Gauss map is the Farey map *accelerated*: $T(x) = F^k(x)$ with $k = \lfloor 1/x \rfloor$ (the k additive steps are done at once).

What is induction

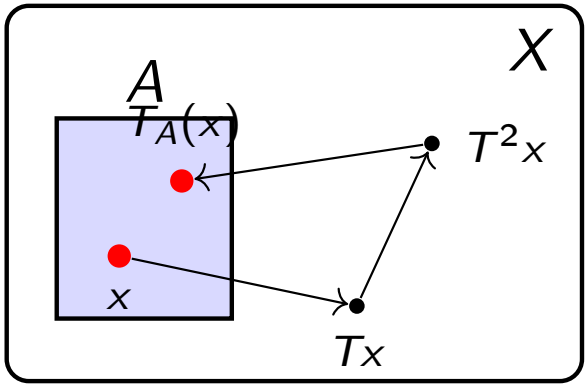
- A dynamical system (X, T) and a subset $A \subset X$ such that every point of A comes back to A.



here $n_x = 3$.

What is induction

- A dynamical system (X, T) and a subset $A \subset X$ such that every point of A comes back to A.
- **First return time**: for $x \in A$, $n_x = \min\{n \geq 1 : \ T^n x \in A\}$.

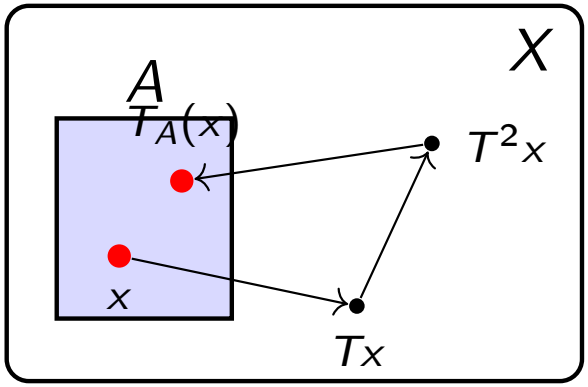


here $n_x = 3$.

What is induction

- A dynamical system (X, T) and a subset $A \subset X$ such that every point of A comes back to A.
- **First return time**: for $x \in A$, $n_x = \min\{n \geq 1 : \ T^n x \in A\}$.
- **First return map** (induced map):

$$T_A : A \to A, \qquad T_A(x) = T^{n_x}(x).$$



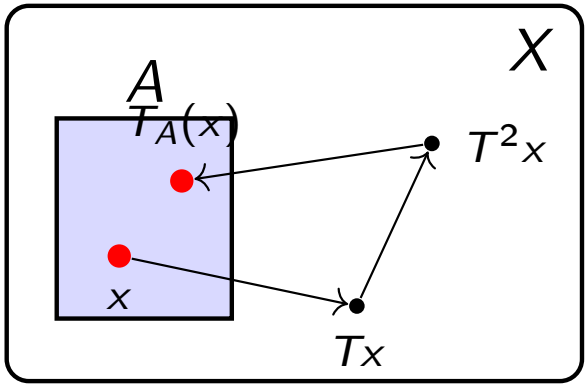
here $n_x = 3$.

What is induction

- A dynamical system (X, T) and a subset $A \subset X$ such that every point of A comes back to A.
- **First return time**: for $x \in A$, $n_x = \min\{n \geq 1 : \ T^n x \in A\}$.
- **First return map** (induced map):

$$T_A : A \to A, \qquad T_A(x) = T^{n_x}(x).$$

- The orbits of T_A are the pieces of the T-orbits that lie in A: T_A is the system T *observed only when it visits* A.



here $n_x = 3$.

Induction: what is preserved

- If μ is a T-invariant probability, almost every point of A returns to A (Poincaré recurrence), and $\mu(\cdot|A)$ is T_A-invariant.

Induction: what is preserved

- If μ is a T-invariant probability, almost every point of A returns to A (Poincaré recurrence), and $\mu(\cdot|A)$ is T_A-invariant.
- **Kac's lemma** (T ergodic): the mean return time is $\int_A n_x \, d\mu(\cdot|A) = \frac{1}{\mu(A)}$.

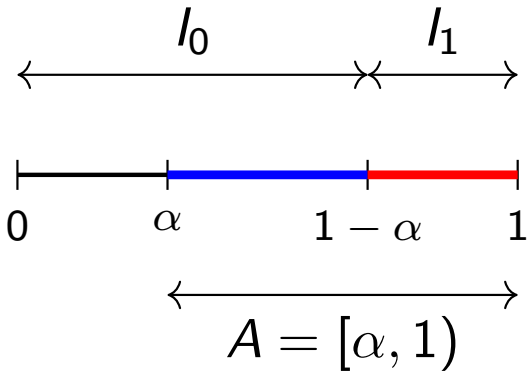
Induction: what is preserved

- If μ is a T-invariant probability, almost every point of A returns to A (Poincaré recurrence), and $\mu(\cdot|A)$ is T_A-invariant.
- **Kac's lemma** (T ergodic): the mean return time is $\int_A n_x \, d\mu(\cdot|A) = \frac{1}{\mu(A)}$.
- If T is ergodic, so is T_A.

Induction: what is preserved

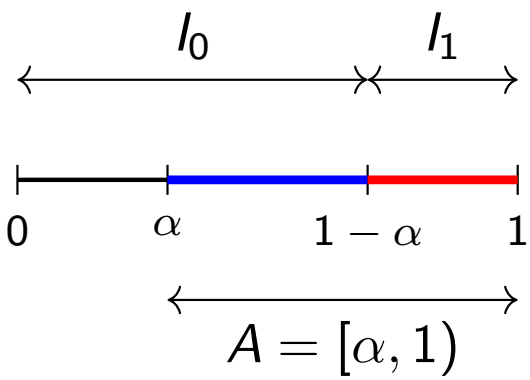
- If μ is a T-invariant probability, almost every point of A returns to A (Poincaré recurrence), and $\mu(\cdot|A)$ is T_A-invariant.
- **Kac's lemma** (T ergodic): the mean return time is $\int_A n_x \, d\mu(\cdot|A) = \frac{1}{\mu(A)}$.
- If T is ergodic, so is T_A.
- T is recovered from T_A and the return time n_x (a tower above A): induction loses nothing, it only changes the scale and the clock.

Induction of a rotation on the image of the longest coding interval



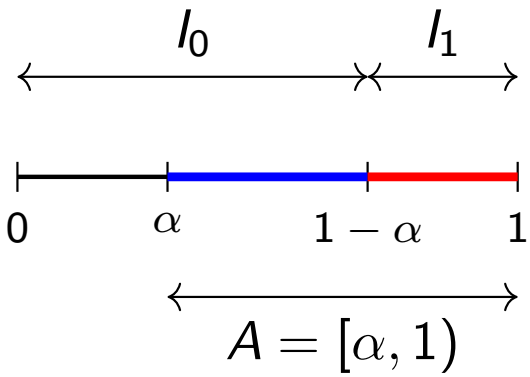
- $R_\alpha : x \mapsto x + \alpha \bmod 1$ is continuous on the **continuity intervals** $I_0 = [0, 1-\alpha)$, $I_1 = [1-\alpha, 1)$ (the coding).

Induction of a rotation on the image of the longest coding interval



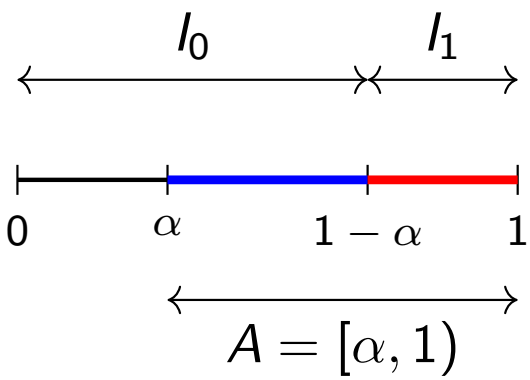
- $R_\alpha : x \mapsto x + \alpha \bmod 1$ is continuous on the **continuity intervals** $I_0 = [0, 1-\alpha)$, $I_1 = [1-\alpha, 1)$ (the coding).
- The **longest** one gives the type: type 0 if I_0 is longer ($\alpha < \frac{1}{2}$), type 1 otherwise.

Induction of a rotation on the image of the longest coding interval



- $R_\alpha : x \mapsto x + \alpha \bmod 1$ is continuous on the **continuity intervals** $I_0 = [0, 1-\alpha)$, $I_1 = [1-\alpha, 1)$ (the coding).
- The **longest** one gives the type: type 0 if I_0 is longer ($\alpha < \frac{1}{2}$), type 1 otherwise.
- Type 0: induce on $A = R_\alpha(I_0) = [\alpha, 1)$, the points whose previous letter is 0.

Induction of a rotation on the image of the longest coding interval



- $R_\alpha : x \mapsto x + \alpha \bmod 1$ is continuous on the **continuity intervals** $I_0 = [0, 1-\alpha)$, $I_1 = [1-\alpha, 1)$ (the coding).
- The **longest** one gives the type: type 0 if I_0 is longer ($\alpha < \frac{1}{2}$), type 1 otherwise.
- Type 0: induce on $A = R_\alpha(I_0) = [\alpha, 1)$, the points whose previous letter is 0.
- Return time 1 on $[\alpha, 1-\alpha)$ (word 0), time 2 on $[1-\alpha, 1)$ (word 10). The induced map is $x \mapsto x + \alpha \bmod (1-\alpha)$.

Induction and the discrete recoding

- After rescaling A to length 1, the induced map is the rotation of angle $\alpha' = \dfrac{\alpha}{1-\alpha}$ (type 0). Its coding is the recoded sequence v, with $u = \sigma_0(v)$: the words 0 and 10 are the blocks of σ_0.

Induction and the discrete recoding

- After rescaling A to length 1, the induced map is the rotation of angle $\alpha' = \dfrac{\alpha}{1-\alpha}$ (type 0). Its coding is the recoded sequence v, with $u = \sigma_0(v)$: the words 0 and 10 are the blocks of σ_0.
- Type 1: induce on $A = R_\alpha(I_1) = [0, \alpha)$. Return time 1 on $[1-\alpha, \alpha)$ (word 1), return time 2 on $[0, 1-\alpha)$ (word 01): the blocks of σ_1; the new angle is $\dfrac{2\alpha - 1}{\alpha}$.

Induction and the discrete recoding

- After rescaling A to length 1, the induced map is the rotation of angle $\alpha' = \dfrac{\alpha}{1-\alpha}$ (type 0). Its coding is the recoded sequence v, with $u = \sigma_0(v)$: the words 0 and 10 are the blocks of σ_0.
- Type 1: induce on $A = R_\alpha(I_1) = [0, \alpha)$. Return time 1 on $[1-\alpha, \alpha)$ (word 1), return time 2 on $[0, 1-\alpha)$ (word 01): the blocks of σ_1; the new angle is $\dfrac{2\alpha - 1}{\alpha}$.
- Lengths of the continuity intervals, type 0: $(1-\alpha, \alpha) \mapsto \dfrac{(1-2\alpha,\ \alpha)}{1-\alpha}$. Inducing on the image of the longest interval **is** the additive Euclid step: we subtract the shorter interval from the longer.

Induction and the discrete recoding

- After rescaling A to length 1, the induced map is the rotation of angle $\alpha' = \dfrac{\alpha}{1-\alpha}$ (type 0). Its coding is the recoded sequence v, with $u = \sigma_0(v)$: the words 0 and 10 are the blocks of σ_0.
- Type 1: induce on $A = R_\alpha(I_1) = [0, \alpha)$. Return time 1 on $[1-\alpha, \alpha)$ (word 1), return time 2 on $[0, 1-\alpha)$ (word 01): the blocks of σ_1; the new angle is $\dfrac{2\alpha - 1}{\alpha}$.
- Lengths of the continuity intervals, type 0: $(1-\alpha, \alpha) \mapsto \dfrac{(1-2\alpha,\ \alpha)}{1-\alpha}$. Inducing on the image of the longest interval **is** the additive Euclid step: we subtract the shorter interval from the longer.
- Induction on the image of the *shortest* interval, $A = [0, \alpha)$, does several steps at once: multiplicative version, i.e. the Gauss map.

An operation on a family of systems

- The rotations $\{R_\alpha\}$ form a family. Induction on the image of the longest interval, *followed by a rescaling* of A to length 1, sends R_α to $R_{\alpha'}$: an operation $\mathcal{Z}$ on the family, acting on the parameter by $\alpha \mapsto \dfrac{\alpha}{1-\alpha}$ or $\dfrac{2\alpha-1}{\alpha}$.

An operation on a family of systems

- The rotations $\{R_\alpha\}$ form a family. Induction on the image of the longest interval, *followed by a rescaling* of A to length 1, sends R_α to $R_{\alpha'}$: an operation $\mathcal{Z}$ on the family, acting on the parameter by $\alpha \mapsto \dfrac{\alpha}{1-\alpha}$ or $\dfrac{2\alpha-1}{\alpha}$.
- It acts as a **zoom**: we look at the same dynamics through a microscope, with smaller scale (length of A) and faster clock.

An operation on a family of systems

- The rotations $\{R_\alpha\}$ form a family. Induction on the image of the longest interval, *followed by a rescaling* of A to length 1, sends R_α to $R_{\alpha'}$: an operation $\mathcal{Z}$ on the family, acting on the parameter by $\alpha \mapsto \dfrac{\alpha}{1-\alpha}$ or $\dfrac{2\alpha-1}{\alpha}$.
- It acts as a **zoom**: we look at the same dynamics through a microscope, with smaller scale (length of A) and faster clock.
- Iterating $\mathcal{Z}$ gives an infinite sequence of zooms, each of type 0 or 1. The numbers of consecutive zooms of the same type are the exponents a_n of Lecture 2, and $\alpha = [0; a_0 + 1, a_1, a_2, \ldots]$.

An operation on a family of systems

- The rotations $\{R_\alpha\}$ form a family. Induction on the image of the longest interval, *followed by a rescaling* of A to length 1, sends R_α to $R_{\alpha'}$: an operation $\mathcal{Z}$ on the family, acting on the parameter by $\alpha \mapsto \dfrac{\alpha}{1-\alpha}$ or $\dfrac{2\alpha-1}{\alpha}$.
- It acts as a **zoom**: we look at the same dynamics through a microscope, with smaller scale (length of A) and faster clock.
- Iterating $\mathcal{Z}$ gives an infinite sequence of zooms, each of type 0 or 1. The numbers of consecutive zooms of the same type are the exponents a_n of Lecture 2, and $\alpha = [0; a_0 + 1, a_1, a_2, \ldots]$.
- This is **renormalization**: the family is stable under the zoom, which acts on the parameter as a one-dimensional dynamical system (Farey, Gauss). The same idea works for interval exchanges (Rauzy–Veech induction).

The shift on sequences of integers

- $\Sigma = (\mathbb{N}^*)^{\mathbb{N}^*}$: sequences $(a_1, a_2, \ldots)$ of positive integers, with the product topology ($\mathbb{N}^*$ discrete), and the **shift** $\sigma(a_1, a_2, a_3, \ldots) = (a_2, a_3, \ldots)$.

The shift on sequences of integers

- $\Sigma = (\mathbb{N}^*)^{\mathbb{N}^*}$: sequences $(a_1, a_2, \ldots)$ of positive integers, with the product topology ($\mathbb{N}^*$ discrete), and the **shift** $\sigma(a_1, a_2, a_3, \ldots) = (a_2, a_3, \ldots)$.
- **Coding**: $\pi : (0,1) \setminus \mathbb{Q} \to \Sigma$, $\pi(x) = (a_1, a_2, \ldots)$, is a bijection and $\pi \circ T = \sigma \circ \pi$: the Gauss map is conjugate to the full shift on a countable alphabet.

The shift on sequences of integers

- $\Sigma = (\mathbb{N}^*)^{\mathbb{N}^*}$: sequences $(a_1, a_2, \ldots)$ of positive integers, with the product topology ($\mathbb{N}^*$ discrete), and the **shift** $\sigma(a_1, a_2, a_3, \ldots) = (a_2, a_3, \ldots)$.
- **Coding**: $\pi : (0,1) \setminus \mathbb{Q} \to \Sigma$, $\pi(x) = (a_1, a_2, \ldots)$, is a bijection and $\pi \circ T = \sigma \circ \pi$: the Gauss map is conjugate to the full shift on a countable alphabet.
- One-to-one: $x = \lim p_n/q_n$ is determined by its digits. Onto: any (a_n) gives a sequence of nested intervals shrinking to a point.

The shift on sequences of integers

- $\Sigma = (\mathbb{N}^*)^{\mathbb{N}^*}$: sequences $(a_1, a_2, \ldots)$ of positive integers, with the product topology ($\mathbb{N}^*$ discrete), and the **shift** $\sigma(a_1, a_2, a_3, \ldots) = (a_2, a_3, \ldots)$.
- **Coding**: $\pi : (0,1) \setminus \mathbb{Q} \to \Sigma$, $\pi(x) = (a_1, a_2, \ldots)$, is a bijection and $\pi \circ T = \sigma \circ \pi$: the Gauss map is conjugate to the full shift on a countable alphabet.
- One-to-one: $x = \lim p_n/q_n$ is determined by its digits. Onto: any (a_n) gives a sequence of nested intervals shrinking to a point.
- **Cylinders**: $[a_1 \ldots a_n] = \{x : \ a_1(x) = a_1, \ldots, a_n(x) = a_n\}$ is the interval with endpoints $\frac{p_n}{q_n}$ and $\frac{p_n+p_{n-1}}{q_n+q_{n-1}}$, of length $\dfrac{1}{q_n(q_n+q_{n-1})}$.

Periodic orbits

- Periodic orbits of $T \leftrightarrow$ periodic sequences $(a_1, \ldots, a_p)^\infty$.
  They are dense in Σ: repeat a finite word.

Periodic orbits

- Periodic orbits of $T \leftrightarrow$ periodic sequences $(a_1, \ldots, a_p)^\infty$. They are dense in Σ: repeat a finite word.
- If $x = [0; \overline{a_1, \ldots, a_p}]$, then $x_p = x$, so $x = \frac{p_p + p_{p-1}x}{q_p + q_{p-1}x}$, i.e. $q_{p-1}x^2 + (q_p - p_{p-1})x - p_p = 0$: x is a **quadratic irrational**.

Periodic orbits

- Periodic orbits of $T \leftrightarrow$ periodic sequences $(a_1, \ldots, a_p)^\infty$. They are dense in Σ: repeat a finite word.
- If $x = [0; \overline{a_1, \ldots, a_p}]$, then $x_p = x$, so $x = \frac{p_p + p_{p-1}x}{q_p + q_{p-1}x}$, i.e. $q_{p-1}x^2 + (q_p - p_{p-1})x - p_p = 0$: x is a **quadratic irrational**.
- **Theorem (Lagrange, 1770)**: the continued fraction of x is *eventually periodic* if and only if x is a quadratic irrational.

Periodic orbits

- Periodic orbits of $T \leftrightarrow$ periodic sequences $(a_1, \ldots, a_p)^\infty$. They are dense in Σ: repeat a finite word.
- If $x = [0; \overline{a_1, \ldots, a_p}]$, then $x_p = x$, so $x = \frac{p_p + p_{p-1}x}{q_p + q_{p-1}x}$, i.e. $q_{p-1}x^2 + (q_p - p_{p-1})x - p_p = 0$: x is a **quadratic irrational**.
- **Theorem (Lagrange, 1770)**: the continued fraction of x is *eventually periodic* if and only if x is a quadratic irrational.
- "$\Rightarrow$": as above. "$\Leftarrow$": the remainders x_n satisfy integer quadratic equations with the *same discriminant* and bounded coefficients: finitely many possibilities, so $x_n = x_m$ for some $n < m$.

Periodic orbits

- Periodic orbits of $T \leftrightarrow$ periodic sequences $(a_1, \ldots, a_p)^\infty$. They are dense in Σ: repeat a finite word.
- If $x = [0; \overline{a_1, \ldots, a_p}]$, then $x_p = x$, so $x = \frac{p_p + p_{p-1}x}{q_p + q_{p-1}x}$, i.e. $q_{p-1}x^2 + (q_p - p_{p-1})x - p_p = 0$: x is a **quadratic irrational**.
- **Theorem (Lagrange, 1770)**: the continued fraction of x is *eventually periodic* if and only if x is a quadratic irrational.
- "$\Rightarrow$": as above. "$\Leftarrow$": the remainders x_n satisfy integer quadratic equations with the *same discriminant* and bounded coefficients: finitely many possibilities, so $x_n = x_m$ for some $n < m$.
- **Theorem (Galois)**: the expansion of $x \in (0, 1)$ is *purely* periodic iff x is a quadratic irrational whose conjugate x' satisfies $x' < -1$. For example $\sqrt{2} - 1 = [0; \overline{2}]$ and $\sqrt{3} - 1 = [0; \overline{1, 2}]$.