Bounded remainder sets and dynamics

V. Berthé

IRIF-CNRS-Paris-France

Equidistribution: Arithmetic, Computational and Probabilistic Aspects
Outline

- From Kronecker sequences and discrepancy to symbolic dynamics
- Topological methods for bounded discrepancy/deviation for frequencies
- From letters to factors
- Gottschalk–Hedlund’s theorem and topological eigenvalues
- Examples
  - Hypercubic billiard
  - Substitutions
Kronecker sequences and discrepancy
Kronecker sequences and discrepancy

Discrepancy Let \((u_n)_n\) be a sequence with values in \([0, 1]\)

\[
\Delta_N = \sup_{\text{interval}} \left| \left\{ \text{Card } 0 \leq n \leq N; u_n \in I \right\} - N\mu(I) \right|
\]
Kronecker sequences and discrepancy

**Discrepancy** Let \((u_n)_n\) be a sequence with values in \([0, 1]\)

\[
\Delta_N = \sup_{I \text{ interval}} \left| \text{Card} \{0 \leq n \leq N; u_n \in I\} - N \mu(I) \right|
\]

**Kronecker sequence** Let \(\alpha = (\alpha_1, \ldots, \alpha_d) \in [0, 1]^d\) with \(1, \alpha_1, \ldots, \alpha_d\) \(\mathbb{Q}\)-linearly independent. Consider the sequence in \([0, 1]^d\)

\[
(\{n\alpha_1\}, \ldots, \{n\alpha_d\})_n
\]

associated with the translation over \(\mathbb{T}^d = (\mathbb{R}/\mathbb{Z})^d\)

\[
R_\alpha : \mathbb{T}^d \mapsto \mathbb{T}^d, \ x \mapsto x + \alpha
\]

**Discrepancy**

\[
\Delta_N = \sup_{B \text{ box}} \left| \text{Card} \{0 \leq n < N; R_\alpha^n(0) \in B\} - N \cdot \mu(B) \right|
\]
Kronecker sequences and discrepancy

Discrepancy Let \((u_n)_n\) be a sequence with values in \([0, 1]\)

\[
\Delta_N = \sup_{I \text{ interval}} \left| \left\{ \text{Card} \{0 \leq n \leq N; u_n \in I\} - N \mu(I) \right\} \right|
\]

Kronecker sequence Let \(\alpha = (\alpha_1, \ldots, \alpha_d) \in [0, 1]^d\) with
\(1, \alpha_1, \ldots, \alpha_d \in \mathbb{Q}\)-linearly independent. Consider the sequence in \([0, 1]^d\)

\[(\{n\alpha_1\}, \ldots, \{n\alpha_d\})_n\]

associated with the translation over \(\mathbb{T}^d = (\mathbb{R}/\mathbb{Z})^d\)

\[R_\alpha : \mathbb{T}^d \mapsto \mathbb{T}^d, \ x \mapsto x + \alpha\]

Discrepancy

\[\Delta_N = \sup_{B \text{ box}} \left| \text{Card} \{0 \leq n < N; R_\alpha^n(0) \in B\} - N \cdot \mu(B) \right|
\]

Bounded remainder set A measurable set \(X\) for which there exists \(C > 0\) s.t. for all \(N\)

\[|\text{Card}\{0 \leq n \leq N; R_\alpha^n(0) \in X\} - N \mu(X)| \leq C\]
Discrepancy

\[ \Delta_N = \sup_{B \text{ box}} |\text{Card} \{0 \leq n < N; R^n_\alpha(0) \in B\} - N \cdot \mu(B)| \]

**Theorem** \( d = 1 \) [Behnke] \( O(\log N) \) if and only if the sequence of the Cesàro means of the partial quotients of \( \alpha \) is bounded
Discrepancy

$$\Delta_N = \sup_{B \text{ box}} |\text{Card} \{0 \leq n < N; R^n_\alpha(0) \in B\} - N \cdot \mu(B)|$$

Theorem $d = 1$ [Behnke]  $O(\log N)$ if and only if the sequence of the Cesàro means of the partial quotients of $\alpha$ is bounded

Theorem [Khintchine][Beck]
For all positive increasing $\varphi$, for a.e. $\alpha$

$$\Delta_N(\alpha) = O((\log N)^d \cdot \varphi(\log \log N)) \quad \text{iff} \quad \sum \frac{1}{\varphi(N)} = +\infty$$

$\sim$ between $(\log N)^d \log \log N$ and $(\log N)^d (\log \log N)^{1+\varepsilon}$
Bounded remainder sets for toral translations

Let \( \alpha = (\alpha_1, \ldots, \alpha_d) \in [0, 1]^d \), \( R_{\alpha}: \mathbb{T}^d \to \mathbb{T}^d, \ x \mapsto x + \alpha \)

Bounded remainder set A measurable set \( X \) for which there exists \( C > 0 \) s.t. for all \( N \)

\[
|\text{Card}\{0 \leq n \leq N; R^n_{\alpha}(0) \in X\} - N\mu(X)| \leq C
\]

Sets with bounded ergodic deviations

[Kesten’66] \( d = 1 \) Intervals that are bounded remainder sets are the intervals with length in \( \mathbb{Z} + \alpha\mathbb{Z} \)

[Liardet’87] \( d \geq 2 \) There are no nontrivial boxes that are bounded remainder sets

[Grepstad-Lev, Haynes-Kelly-Koivusalo] Any parallelepiped in \( \mathbb{R}^d \) spanned by vectors \( v_1, \ldots, v_d \) belonging to \( \mathbb{Z} \alpha + \mathbb{Z}^d \) is a bounded remainder set for the translation by \( (\alpha_1, \ldots, \alpha_d) \) on \( \mathbb{T}^d \), with \( 1, \alpha_1, \ldots, \alpha_d \) linearly independent.
A symbolic approach

Consider the translation over \( T^d = (\mathbb{R}/\mathbb{Z})^d \)

\[
R_\alpha : T^d \mapsto T^d, \ x \mapsto x + \alpha \quad \text{with} \ \alpha = (\alpha_1, \ldots, \alpha_d)
\]

We consider a partition \( \{X_1, \ldots, X_k\} \) of \( T^d \)

\[
T^d = \bigcup_{1 \leq i \leq k} X_i, \quad \mu(X_i \cap X_j) = 0, \quad \text{for all} \ i \neq j
\]

We code the trajectory of \( x \) under the action of \( R_\alpha : x \mapsto x + \alpha \) as follows

\[
x \sim (u_n)_n \in \{1, 2, \ldots, k\}^\mathbb{N}
\]

\[
u_n = i \quad \text{if and only if} \quad R^n_\alpha(x) = x + n\alpha \in X_i
\]
A symbolic approach

Consider the translation over $\mathbb{T}^d = (\mathbb{R}/\mathbb{Z})^d$

$$R_\alpha: \mathbb{T}^d \mapsto \mathbb{T}^d, \ x \mapsto x + \alpha \quad \text{with} \ \alpha = (\alpha_1, \ldots, \alpha_d)$$

We consider a partition $\{X_1, \cdots, X_k\}$ of $\mathbb{T}^d$

$$\mathbb{T}^d = \bigcup_{1 \leq i \leq k} X_i, \quad \mu(X_i \cap X_j) = 0, \ \text{for all} \ i \neq j$$

We code the trajectory of $x$ under the action of $R_\alpha: x \mapsto x + \alpha$ as follows

$$x \sim (u_n)_n \in \{1, 2, \ldots, k\}^\mathbb{N}$$

$$u_n = i \quad \text{if and only if} \quad R^n_\alpha(x) = x + n\alpha \in X_i$$

Questions  Which information on $R_\alpha$ can we get from the combinatorial properties of the sequence $(u_n)$? What is a good coding?
Symbolic dynamics

- **1898, Hadamard**: Geodesic flows on surfaces of negative curvature
- **1912, Thue**: Prouhet-Thue-Morse substitution
  \[ \sigma : a \mapsto ab, \ b \mapsto ba \]
- **1921, Morse**: Symbolic representation of geodesics on a surface with negative curvature. Recurrent geodesics
Symbolic dynamics

- **1898, Hadamard**: Geodesic flows on surfaces of negative curvature
- **1912, Thue**: Prouhet-Thue-Morse substitution
  \[ \sigma : a \mapsto ab, \ b \mapsto ba \]
- **1921, Morse**: Symbolic representation of geodesics on a surface with negative curvature. Recurrent geodesics

From geometric dynamical systems to symbolic dynamical systems and backwards

- Given a geometric system, can one find a good partition?
- And vice-versa?
Symbolic dynamical systems

Let $u \in \mathcal{A}^\mathbb{Z}$ be an infinite word

$$u = \cdots abaababaababaababaababaababaababaababaabababaabaababaababaababaab \cdots$$

$aa$ is a factor, $bb$ is not a factor

Let $\mathcal{L}_u$ be the set of factors of $u$: $\mathcal{L}_u$ is the language of $u$
Symbolic dynamical systems

Let $u \in \mathcal{A}^\mathbb{Z}$ be an infinite word

$$u = \cdots abaababaabaababaababaababaababaababaababaababaababaab \cdots$$

$aa$ is a factor, $bb$ is not a factor

Let $\mathcal{L}_u$ be the set of factors of $u$: $\mathcal{L}_u$ is the language of $u$

$\mathcal{A}^\mathbb{Z}$ is a compact metric space $d(u, \nu) = 2^{-\max\{k : u_{[-k \ldots k]} = \nu_{[-k \ldots k]}\}}$

The shift $T$ acts on $\mathcal{A}^\mathbb{Z}$ as $T((u_n)_n) = (u_{n+1})_n$

The symbolic dynamical system generated by $u$ is $(X_u, T)$ with

$$X_u := \overline{\{T^n(u) ; \ n \in \mathbb{Z}\}} = \{\nu \in \mathcal{A}^\mathbb{Z} ; \mathcal{L}_\nu \subset \mathcal{L}_u\} \subset \mathcal{A}^\mathbb{Z}$$

A subshift $(X, T)$ is a closed shift-invariant subset of $\mathcal{A}^\mathbb{Z}$
Ergodic theorem

Among the first $N$ points of the orbit of $x$, how many of them enter $B$? How often do they visit $B$?

$$\lim_{N \to \infty} \frac{1}{N} \sum_{0 \leq n < N} 1_{B}(T^n x) = \mu(P) \text{ a.e.}$$
Ergodic theorem

Among the first $N$ points of the orbit of $x$, how many of them enter $B$?

How often do they visit $B$?
Let $1_B$ be the characteristic function of $B$

Among the first $N$ points of the orbit of $x$, how many of them enter $B$?

$$\sum_{0 \leq n < N} 1_B(T^n x)$$

How often do they visit $B$?

$$\lim_{N \to \infty} \frac{1}{N} \sum_{0 \leq n < N} 1_B(T^n x)$$

$$\lim_{N \to \infty} \frac{1}{N} \sum_{0 \leq n < N} 1_B(T^n x) = \mu(P) \text{ a.e. } x$$
Ergodic theorem

We are given a dynamical system \((X, T, \mathcal{B}, \mu)\) with \(T : X \to X\)

- **Average time values**: one particle over the long term
- **Average space values**: all particles at a particular instant

Ergodicity

\[ \mu(B) = \mu(T^{-1}B) \]  \(T\)-invariance

\[ T^{-1}B = B \implies \mu(B) = 0 \text{ or } 1 \text{ ergodicity} \]

Ergodic theorem  \(\text{space average} = \text{time average}\)

\[ f \in L_1(\mu) \implies \lim_{N \to \infty} \frac{1}{N} \sum_{0 \leq n < N} f(T^n x) = \int f \, d\mu \quad \text{a.e. } x \]
Examples of dynamical systems

- Translation on the torus \( R_\alpha : x \mapsto \alpha + x \mod 1 \)
- Symbolic dynamical systems \((X, T)\) where \( T \) is the shift
- Beta-maps \( T : [0, 1] \to [0, 1], \ x \mapsto \{ \beta x \} \)
- Continued fractions \( T : [0, 1] \to [0, 1], \ x \mapsto \{ 1/x \} \)
Examples of dynamical systems

- Translation on the torus $\mathcal{R}_\alpha : x \mapsto \alpha + x \mod 1$ zero entropy
- Symbolic dynamical systems $(X, T)$ where $T$ is the shift
- Beta-maps $T : [0, 1] \rightarrow [0, 1], \ x \mapsto \{\beta x\}$ positive entropy
- Continued fractions $T : [0, 1] \rightarrow [0, 1], \ x \mapsto \{1/x\}$ positive entropy
Sturmian dynamics

Sturmian words are codings of Kronecker translations on $\mathbb{R}/\mathbb{Z}$

$$R_\alpha : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z}, \ x \mapsto x + \alpha \mod 1$$

according to the finite partition

$$\{ l_0 = [0, 1 - \alpha[, \ l_1 = [1 - \alpha, 1] \}$$
Sturmian dynamics

Sturmian words are codings of Kronecker translations on $\mathbb{R}/\mathbb{Z}$

$$R_\alpha : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}, \ x \mapsto x + \alpha \mod 1$$

according to the finite partition

$$\{ l_0 = [0, 1 - \alpha[, \ l_1 = [1 - \alpha, 1]\}$$

This yields a measure-theoretic isomorphism

$$\begin{array}{ccc}
\mathbb{R}/\mathbb{Z} & \xrightarrow{R_\alpha} & \mathbb{R}/\mathbb{Z} \\
\uparrow & & \uparrow \\
X_\alpha & \xrightarrow{T} & X_\alpha
\end{array}$$

where $T$ is the shift and $X_\alpha \subset \{0, 1\}^\mathbb{N}$
Strong uniformity and ergodicity [Beck]

For every interval $I$

$$\lim_{N \to \infty} \frac{1}{N} \sum_{0 \leq n \leq N-1} 1_I(n\alpha) = \mu(I)$$

for every irrational $\alpha$

Khinchin’s conjecture’23 For any Lebesgue measurable set $S \subset [0, 1]$, the sequence $(n\alpha)_n$ modulo 1 is uniformly distributed with respect to $S$ for almost every $\alpha$
Strong uniformity and ergodicity [Beck]

For every interval $I$

$$\lim \frac{1}{N} \sum_{0 \leq n \leq N-1} 1_I(n\alpha) = \mu(I) \quad \text{for every irrational } \alpha$$

Khinchin’s conjecture’23 For any Lebesgue measurable set $S \subset [0, 1]$, the sequence $(n\alpha)_n$ modulo 1 is uniformly distributed with respect to $S$ for almost every $\alpha$

$$\lim \frac{1}{N} \sum_{0 \leq n \leq N-1} 1_S(n\alpha) = \mu(I) \quad \text{for a.e. } \alpha$$

It was disproved by [Marstrand’70]

There exists an open set $S$ with $\mu(S) < 1$ such that

$$\lim \sup \frac{1}{N} \sum_{0 \leq n \leq N-1} 1_S(n\alpha) = 1 \quad \text{for every } \alpha$$
Strong uniformity and ergodicity [Beck]

For every interval $I$

$$\lim_{N \to \infty} \frac{1}{N} \sum_{0 \leq n \leq N-1} 1_I(n \alpha) = \mu(I) \quad \text{for every irrational } \alpha$$

Khinchin’s conjecture'23 For any Lebesgue measurable set $S \subset [0,1]$, the sequence $(n \alpha)_n$ modulo 1 is uniformly distributed with respect to $S$ for almost every $\alpha$ Disproved by [Marstrand’70]

Positive results

$$\lim_{N \to \infty} \frac{1}{N} \sum_{0 \leq n \leq N-1} 1_I(n \alpha + \beta) = \mu(I) \text{ for a.e. } \beta$$

Ergodicity of $R_\alpha: x \mapsto x + \alpha$ on $\mathbb{T}$

Raikov For every Lebesgue measurable set $S \subset [0,1]$

$$\lim_{N \to \infty} \frac{1}{N} \sum_{0 \leq n \leq N-1} 1_S(2^n \alpha) = \mu(I) \text{ for a.e. } \alpha$$

Ergodicity of $T_2: x \mapsto 2x$ on $\mathbb{T}$
Symbolic discrepancy
Frequencies

Take an infinite word \( u = (u_n)_n \) with values in a finite alphabet \( \mathcal{A} \).

The frequency \( f_i \) of the letter \( i \) in \( u \) is defined as the following limit, if it exists

\[
f_i = \lim_{n \to \infty} \frac{|u_0 \cdots u_{N-1}|_i}{N}
\]
Frequencies

Take an infinite word \( u = (u_n)_n \) with values in a finite alphabet \( \mathcal{A} \).

The frequency \( f_i \) of the letter \( i \) in \( u \) is defined as the following limit, if it exists:

\[
 f_i = \lim_{n \to \infty} \frac{|u_0 \cdots u_{N-1}|_i}{N}
\]

One defines similar notions for factors.

Frequencies do not always exist.

\[
u = abaabbaaaabbbbaaaaaaaabbbbbb
\]
Frequencies and unique ergodicity

The frequency $f_i$ of a letter $i$ in $u$ is defined as the following limit, if it exists

$$f_i = \lim_{n \to \infty} \frac{|u_0 \cdots u_{N-1}|_i}{N}$$
Frequencies and unique ergodicity

The frequency $f_i$ of a letter $i$ in $u$ is defined as the following limit, if it exists

$$f_i = \lim_{n \to \infty} \frac{|u_0 \cdots u_{N-1}|_i}{N}$$

One can also consider

$$\lim_{n \to \infty} \frac{|u_k \cdots u_{k+N-1}|_i}{N}$$

This corresponds to well distribution.
If the convergence is uniform with respect to $k$, one says that $u$ has uniform letter frequencies.

One defines similar notions for factors.
Frequencies and unique ergodicity

The frequency \( f_i \) of a letter \( i \) in \( u \) is defined as the following limit, if it exists

\[
    f_i = \lim_{n \to \infty} \frac{|u_0 \cdots u_{N-1}|_i}{N}
\]

One can also consider

\[
    \lim_{n \to \infty} \frac{|u_k \cdots u_{k+N-1}|_i}{N}
\]

This corresponds to well distribution.
If the convergence is uniform with respect to \( k \), one says that \( u \) has uniform letter frequencies.

One defines similar notions for factors.

The symbolic shift \((X_u, T)\) is said to be uniquely ergodic if \( u \) has uniform factor frequency for every factor.

Equivalently, there exists a unique shift-invariant probability measure on the symbolic shift \((X_u, T)\).
Factor complexity and frequencies

The factor complexity of an infinite word $u$ counts the number of factors of a given length in $u$.

Both notions have nothing to do a priori, but...

Factor complexity has to do with unique ergodicity in some cases.

- A word is **uniformly recurrent** if every factor occurs with bounded gaps.

  A uniformly recurrent word $u$ such that $\limsup p_u(n)/n < 3$ is uniquely ergodic [Boshernitzan]. 😊

- There exist regular (idoc) interval exchanges on a four-letter alphabet ($p(n) = 3n + 1$ for all $n$) that are not uniquely ergodic [Keane]. 😞
Frequencies and symbolic discrepancy

- **Frequency** $\mu_v$ of a finite word $v$ in $u \in A^\mathbb{N}$

$$\mu_v = \lim_{n \to +\infty} \frac{|u_0 \cdots u_{N-1}|_v}{N}$$

$|u_0 \cdots u_{N-1}|_v =$ number of occurrences of $v$ in $u_0 \cdots u_{N-1}$

- **Symbolic discrepancy**

$$\Delta_u(v, N) = ||u_0 u_1 \cdots u_{N-1}|_v - N \cdot \mu_v|$$

$$\Delta_X(v, N) = \sup_{w \in \mathcal{L}_X(N)} ||w|_v - N \cdot \mu_v|$$ for $X$ subshift

- **Remark** Let $(X, T)$ be a minimal subshift (every closed shift-invariant subset of $X$ either empty or $X$). Every $x$ in $X$ admits frequencies for every factor iff $(X, T)$ is uniquely ergodic [Oxtoby]
Examples

- Let $X_\sigma$ be the Fibonacci shift generated by
  
  $\sigma : 0 \mapsto 01, 1 \mapsto 0$

Let $\mathcal{L}_{X_\sigma}$ be the set of factors of $X_\sigma$

For $\nu \in \mathcal{L}_{X_\sigma}$ $\Delta_X(\nu, N)$ is bounded, i.e., there exist $C > 0$ such that for any factor $w \in \mathcal{L}_X$

$$||w|_\nu - \mu_w||_w \leq C$$
Examples

• Let $X_\sigma$ be the Fibonacci shift generated by

$$\sigma : 0 \mapsto 01, \ 1 \mapsto 0$$

Let $\mathcal{L}_{X_\sigma}$ be the set of factors of $X_\sigma$
For $v \in \mathcal{L}_{X_\sigma}$ $\Delta_X(v, N)$ is bounded, i.e., there exist $C > 0$ such that for any factor $w \in \mathcal{L}_X$

$$||w|_v - \mu_v|w|| \leq C$$

• Let $X_\sigma$ be the Thue-Morse shift generated by

$$\sigma : 0 \mapsto 01, \ 1 \mapsto 10$$

The symbolic discrepancy is bounded for letters but not for words of length $\geq 2$. 
A combinatorial viewpoint: balancedness

Let \((X, T)\) be a minimal subshift

\((X, T)\) is balanced on \(v \in \mathcal{L}_X\) if there exists \(C > 0\) such that for any pair \((x, y)\) of factors of the same length

\[
|||x|_v - |y|_v| \leq C
\]

\((X, T)\) is balanced on factors if it is balanced on all \(v \in \mathcal{L}_X\)

Balance and frequencies The minimal subshift \((X, T)\) is balanced on the factor \(v\) iff there exist \(C > 0\) and \(\mu_v\) such that for any factor \(w \in \mathcal{L}_X\)

\[
||w|_v - \mu_v|w|| \leq C
\]

Remark Does not require to know the frequency \(\mu_v\)
Discrepancy and coboundaries

The discrepancy of \( \nu \) is bounded for \((X, T)\) iff the ergodic sums for \( f_\nu = 1_{[\nu]} - \mu_\nu \) are bounded

\[
\sum_{n=0}^{N-1} f_\nu(T^n(u)) = |u_0 \cdots u_{N+|\nu|-1}|_\nu - \mu_\nu N
\]

\( f \) is a coboundary iff its ergodic sums are bounded
Discrepancy and coboundaries

The discrepancy of $v$ is bounded for $(X, T)$ iff the ergodic sums for $f_v = 1_{[v]} - \mu_v$ are bounded

$$\sum_{n=0}^{N-1} f_v(T^n(u)) = |u_0 \cdots u_{N+|v|-1}|_v - \mu_v N$$

$f$ is a coboundary iff its ergodic sums are bounded

**Theorem [Gottschalk-Hedlund]** Let $X$ be a compact metric space and $T: X \to X$ be a minimal homeomorphism. Let $f: X \to \mathbb{R}$ be a continuous function. Then $f$ is a coboundary

$$f = g - g \circ T$$

for a continuous function $g$ if and only if there exists $x$ and there exists $C > 0$ such that for all $N$

$$| \sum_{n=0}^{N} f(T^n(x)) | < C$$
Discrepancy and coboundaries

The discrepancy of \( \nu \) is bounded for \((X, T)\) iff the ergodic sums for \( f_\nu = 1_{[\nu]} - \mu_\nu \) are bounded

\[
\sum_{n=0}^{N-1} f_\nu(T^n(u)) = |u_0 \cdots u_{N+|\nu|-1}|_\nu - \mu_\nu N
\]

\( f \) is a coboundary iff its ergodic sums are bounded

**Theorem [Gottschalk-Hedlund]** Let \( X \) be a compact metric space and \( T: X \to X \) be a minimal homeomorphism. Let \( f: X \to \mathbb{R} \) be a continuous function. Then \( f \) is a coboundary

\[
f = g - g \circ T
\]

for a continuous function \( g \) if and only if there exists \( x \) and there exists \( C > 0 \) such that for all \( N \)

\[
|\sum_{n=0}^{N} f(T^n(x))| < C
\]

The discrepancy of \( \nu \) is bounded iff \( f_\nu = 1_{[\nu]} - \mu_\nu \) is a coboundary
Bounded discrepancy and topological eigenvalues

\( \nu \) has bounded discrepancy in \((X, T)\) iff \( f_\nu = 1_\nu - \mu_\nu \) is a coboundary

Take \( f = 1_\nu - \mu_\nu \leadsto f = g - g \circ T \)

\[ \exp^{2i\pi g \circ T} = \exp^{2i\pi \mu_\nu} \exp^{2i\pi g} \]

\( \exp^{2i\pi g} \) is a continuous eigenfunction associated with the eigenvalue \( \exp^{2i\pi \mu_\nu} \leadsto \text{Topological rotation factor} \)

If \( \nu \) has bounded discrepancy in \((X, T)\), then \( \mu_\nu \) is an additive topological eigenvalue
Billiards
From factors to intervals

\[ R_{\alpha} : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}, \quad x \mapsto x + \alpha \mod 1 \]
The factors of $u$ are in one-to-one correspondence with the $n+1$ intervals of $\mathbb{T}$ whose end-points are given by

$$-k\alpha \mod 1, \text{ for } 0 \leq k \leq n$$

$$w \sim I_W = I_{w_1} \cap R_0^{-1}I_{w_2} \cap \cdots R_0^{-n+1}I_{w_n}$$

By uniform distribution of $(k\alpha)_k$ modulo 1, the frequency of a factor $w$ of a Sturmian word is equal to the length of $I_w$. 

From factors to intervals
Sturmian words code trajectories in square billiards according to the parallel sides of the square table that the trajectory hits.
Consider a subshift generated by the coding of a trajectory of a billiard in the hypercube of \( \mathbb{R}^d \) with slope \((\alpha_1, \cdots, \alpha_d)\), with \((\alpha_1, \cdots, \alpha_d)\) linearly independent over \( \mathbb{Q} \). It is minimal and uniquely ergodic. Let \( \mu \) be the invariant measure.
Billiard

Consider a subshift generated by the coding of a trajectory of a billiard in the hypercube of $\mathbb{R}^d$ with slope $(\alpha_1, \cdots, \alpha_d)$, with $(\alpha_1, \cdots, \alpha_d)$ linearly independent over $\mathbb{Q}$. It is minimal and uniquely ergodic. Let $\mu$ be the invariant measure.

Cubic case and beyond $d \geq 3$ [Bedaride-B-Jullien]

Letters have bounded discrepancy but factors of length at least 2 have unbounded discrepancy.

Proof Assume that $w$ has bounded discrepancy. Then, $\mu[w]$ is an additive eigenvector and $\mu[w] \in \langle \alpha_1, \cdots, \alpha_d \rangle$. However, the areas of the zones that correspond to factors of length 2 do not belong to $\langle \alpha_1, \cdots, \alpha_d \rangle$.

In other words, the cohomology is not finitely generated.
Consider a subshift generated by the coding of a trajectory of a billiard in the hypercube of $\mathbb{R}^d$ with slope $(\alpha_1, \cdots, \alpha_d)$, with $(\alpha_1, \cdots, \alpha_d)$ linearly independent over $\mathbb{Q}$. It is minimal and uniquely ergodic. Let $\mu$ be the invariant measure.

**Cubic case and beyond $d \geq 3$ [Bedaride-B-Jullien]**

Letters have bounded discrepancy but factors of length at least 2 have unbounded discrepancy.

**Sturmian case $d = 2$**. All factors have bounded discrepancy.
Bounded remainder sets for toral translations

Let $\alpha = (\alpha_1, \ldots, \alpha_d) \in [0, 1]^d$, $R_\alpha : \mathbb{T}^d \rightarrow \mathbb{T}^d$, $x \mapsto x + \alpha$

[Kešten’66] $d = 1$ Intervals that are bounded remainder sets are the intervals with length in $\mathbb{Z} + \alpha \mathbb{Z}$
Factors have bounded discrepancy for Sturmian words.

[Liardet’87] $d \geq 2$ There are no nontrivial boxes that are bounded remainder sets

[Grepstad-Lev, Haynes-Kelly-Koivusalo] Any parallelepiped in $\mathbb{R}^d$ spanned by vectors $v_1, \cdots, v_d$ belonging to $\mathbb{Z} \alpha + \mathbb{Z}^d$ is a bounded remainder set for the translation by $(\alpha_1, \cdots, \alpha_d)$ on $\mathbb{T}^d$, with $1, \alpha_1, \cdots, \alpha_d$ linerly independent.
Letters have bounded discrepancy for hypercubic billiard words.
Continuous version of Khinchin’s conjecture [Beck]

Let $S \subset [0, 1)^2$ be measurable with $0 < \mu(S) < 1$.

$$T_S(\alpha) = \text{meas}\{t \in [0, T]; (\{t \cos \alpha\}, \{t \sin \alpha\}) \in S\}$$

Theorem [Beck] For every $\varepsilon > 0$

$$T_S(\alpha) = \mu(S) T + o((\log T)^{3+\varepsilon}) \text{ for a.e. } \alpha$$

$\sim$ Superuniformity

- $d = 3 \quad T^{\frac{1}{4}}((\log T)^{3+\varepsilon})$
- $d \geq 4 \quad T^{\frac{1}{2} - \frac{1}{2(d-1)}}((\log T)^{3+\varepsilon})$
Dynamical dimension group of a minimal subshift \((X, T)\)

- Coboundaries \(\beta: C(X, \mathbb{Z}) \to C(X, \mathbb{Z}), \ f \mapsto f \circ T - f\)
- Dimension group \(H(X, T) = C(X, \mathbb{Z})/\beta C(X, \mathbb{Z})\)

**Thue-Morse substitution**

\[
H(X, T) = \mathbb{Z}[1/2]
\]

\[
M = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}
\]

\(1_{aa} - 1/6 \neq 0\) in \(H(X, T)\)

Unbounded discrepancy

**Sturmian subshift**

\[
H(X, T) = \mathbb{Z}^2
\]

\[
M = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}
\]

\(1_{[w]} \in \langle 1_{[0]}, 1_{[1]} \rangle\)

Bounded discrepancy

How to decompose a shift \((X, T)\)?

- Kakutani-Rohlin towers and Bratteli-Vershik maps [Herman-Putnam-Skau]
- Return words [Durand-Host-Skau]
- Desubstitution and \(S\)-adic representations
Discrepancy and substitutions
Let $\sigma$ be a primitive substitution.

**Theorem [Adamczewski]**

- If $\sigma$ is a **Pisot substitution**, then letters have bounded discrepancy in $X_\sigma$.

- Conversely, if letters have bounded discrepancy in $X_\sigma$, then the Perron–Frobenius eigenvalue of $M_\sigma$ is the unique eigenvalue of $M_\sigma$ that is larger than 1 in modulus, and all possible eigenvalues of modulus one of $M_\sigma$ are **roots of unity**.
Two-letter factor substitution

Given a substitution $\sigma$, consider the finite set $\mathcal{L}_\sigma(2)$ as an alphabet and define the two-letter factor substitution $\sigma_2$ on $\mathcal{L}_\sigma(2)$ as follows

for every $u = ab \in \mathcal{L}_\sigma(2)$, $\sigma_2(u)$ is the word over $\mathcal{L}_\sigma(2)$ made of the first $|\sigma(a)|$ factors of length 2 in $\sigma(u)$

If $ab \in \mathcal{L}_\sigma(2)$ with $\sigma(a) = a_0 \cdots a_r$, $\sigma(b) = b_0 \cdots b_s$, then

$$\sigma_2(ab) = (a_0a_1)(a_1a_2) \cdots (a_{r-1}a_r)(a_rb_0)$$

- If the substitution $\sigma$ is primitive, then $\sigma_2$ is also primitive, and $\sigma_2$ has the same Perron–Frobenius eigenvalue as $\sigma$
- Frequencies of factors are provided by the renormalized Perron–Frobenius eigenvector of $M_{\sigma_2}$

[Queffélec]
Let $\sigma$ be a primitive substitution.

**Theorem [Adamczewski]**

- If $\sigma$ (resp. $\sigma_2$) is a **Pisot substitution**, then letters (resp. factors) have bounded discrepancy in $X_\sigma$.

- Conversely, if letters (resp. factors) have bounded discrepancy in $X_\sigma$, then the Perron–Frobenius eigenvalue of $M_\sigma$ (resp. $M_{\sigma_2}$) is the unique eigenvalue of $M_\sigma$ (resp. $M_{\sigma_2}$) that is larger than 1 in modulus, and all possible eigenvalues of modulus one of $M_\sigma$ (resp. $M_{\sigma_2}$) are **roots of unity**.
Bounded discrepancy with 1 as an eigenvalue

Example [Cassaigne-Pytheas Fogg-Minervino]
Take
\[ \sigma: 1 \mapsto 121, 2 \mapsto 32, 3 \mapsto 321 \]

The eigenvalues of the substitution matrix are \( \{1, \frac{3 \pm \sqrt{5}}{2}\} \) and it has bounded discrepancy on factors.

Proof Consider the Sturmian substitution
\[ \tau: 3 \mapsto 30, 0 \mapsto 300 \]

The subshift \((X_\sigma, T)\) is deduced from the Sturmian shift \((X_\tau, T)\) by applying the substitution \( \varphi: 0 \mapsto 21, 3 \mapsto 3 \), which preserves balancedness, and thus bounded discrepancy.
Consider the Thue–Morse substitution

\[ \sigma: 0 \mapsto 01, \ 1 \mapsto 10 \]

- One has \( \mathcal{L}_\sigma(2) = \{00, 01, 10, 11\} \)
- One has \( \sigma(00) = 0101 \) and \( \sigma_2(00) = (01)(10) \)
- One checks that \( \sigma^{(2)}(a) \mapsto bc, \ b \mapsto bd, \ c \mapsto ca, \ d \mapsto cb \), by setting \( a = 00, \ b = 01, \ c = 10, \ d = 11 \)
- The eigenvalues of \( M_\sigma \) are 2 and 0, and the eigenvalues of \( M_{\sigma_2} \) are 0, 1, \(-1\) and 2.
- Letters have bounded discrepancy but no factor of length \( \ell \), with \( \ell \geq 2 \), has bounded discrepancy.
- **Remark** If one applies \( \sigma \) to any infinite word in \( \{0, 1\}^\mathbb{Z} \), one gets an infinite word with bounded discrepancy on letters.
How to detect unbounded discrepancy

- Work modulo coboundaries in \( \mathcal{C}(X, \mathbb{Z}) \) (locally constant functions).
- Find a good representation of \((X, T)\) in Kakutani-Rohlin towers.

Thue-Morse shift

- Kakutani-Rohlin towers
  \[ \mathcal{P}_n = \{ T^j \sigma^n([ab]) : ab \in \mathcal{L}_X(2), 0 \leq j < |\sigma^n(a)| \} \]
- This partition is finer than the partition in \(n\)-cylinders

\[
\begin{array}{c|c|c|c}
00 & 01 & 10 & 11 \\
\hline
\end{array}
\]
How to detect unbounded discrepancy for rational frequencies

- KR towers $\mathcal{P}_n = \{ T^j \sigma^n([ab]) : ab \in \mathcal{L}_X(2), 0 \leq j < |\sigma^n(a)| \}$
How to detect unbounded discrepancy for rational frequencies

- KR towers \( \mathcal{P}_n = \{ T^j \sigma^n([ab]) : ab \in \mathcal{L}_X(2), 0 \leq j < |\sigma^n(a)| \} \)
- Take \( f = 1_v - \mu_v \). There exists \( k \) for which \( f \) is constant in the atoms of \( \mathcal{P}_k \). For \( n \geq k \), let \( \phi_n \in \mathbb{R}\mathcal{L}_\sigma(2) \)

\[
\phi_n(ab) = \sum_{j=0}^{\lfloor |\sigma^n(a)|-1 \rfloor} f \mid_{T^j \sigma^n([ab])} \quad \forall ab \in \mathcal{L}_\sigma(2)
\]
How to detect unbounded discrepancy for rational frequencies

- KR towers $\mathcal{P}_n = \{ T^j \sigma^n([ab]) : ab \in \mathcal{L}_X(2), 0 \leq j < |\sigma^n(a)| \}$
- Take $f = 1_v - \mu_v$. There exists $k$ for which $f$ is constant in the atoms of $\mathcal{P}_k$. For $n \geq k$, let $\phi_n \in \mathbb{R}^{\mathcal{L}_\sigma(2)}$

$$\phi_n(ab) = \sum_{j=0}^{|\sigma^n(a)|-1} f \mid T^j \sigma^n([ab]) \quad \forall ab \in \mathcal{L}_\sigma(2)$$

- If $f \in C(X, \mathbb{Z})$ is a coboundary, then it is the coboundary of some $h \in C(X, \mathbb{Z}) \sim$ locally constant. Take $\mu_v$ rational.
- Define $R_n(X) = \{ \phi : \mathcal{L}_X(n) \rightarrow \mathbb{R} \}$ and

$$\beta : R_1(X) \rightarrow R_2(X); \varphi \mapsto (\beta \varphi)(ab) = \varphi(b) - \varphi(a) \quad \forall ab \in \mathcal{L}_X(2)$$
How to detect unbounded discrepancy for rational frequencies

- **KR towers** \( \mathcal{P}_n = \{ T^j \sigma^n([ab]) : ab \in \mathcal{L}_X(2), 0 \leq j < |\sigma^n(a)| \} \)

- Take \( f = 1_v - \mu_v \). There exists \( k \) for which \( f \) is constant in the atoms of \( \mathcal{P}_k \). For \( n \geq k \), let \( \phi_n \in \mathbb{R}^{\mathcal{L}_\sigma(2)} \)

\[
\phi_n(ab) = \sum_{j=0}^{\left| \sigma^n(a) \right|-1} f \big| T^j \sigma^n([ab]) \quad \forall ab \in \mathcal{L}_\sigma(2)
\]

- If \( f \in C(X, \mathbb{Z}) \) is a coboundary, then it is the coboundary of some \( h \in C(X, \mathbb{Z}) \sim \) locally constant. Take \( \mu_v \) rational.

- Define \( R_n(X) = \{ \phi : \mathcal{L}_X(n) \to \mathbb{R} \} \) and

\[
\begin{align*}
\beta : \underbrace{R_1(X)}_{\mathbb{R}^A} &\to \underbrace{R_2(X)}_{\mathbb{R}^{\mathcal{L}_X(2)}}; \varphi \mapsto (\beta \varphi)(ab) = \varphi(b) - \varphi(a) \quad \forall ab \in \mathcal{L}_X(2) \\
\end{align*}
\]

**Theorem** If \( f \) is a coboundary, then \( \phi_n \in \beta(R_1(X_\sigma)) \) for \( n \) large enough [Host, Durand-Host-Perrin]
How to detect unbounded discrepancy for rational frequencies

- KR towers \( \mathcal{P}_n = \{ T^j \sigma^n([ab]) : ab \in \mathcal{L}_X(2), 0 \leq j < |\sigma^n(a)| \} \)
- Take \( f = 1_v - \mu_v \). There exists \( k \) for which \( f \) is constant in the atoms of \( \mathcal{P}_k \). For \( n \geq k \), let \( \phi_n \in \mathbb{R}^{\mathcal{L}_\sigma(2)} \)
  \[ \phi_n(ab) = \sum_{j=0}^{\sigma^n(a)-1} f \big| T^j \sigma^n([ab]) \quad \forall ab \in \mathcal{L}_\sigma(2) \]

- If \( f \in C(X, \mathbb{Z}) \) is a coboundary, then it is the coboundary of some \( h \in C(X, \mathbb{Z}) \leadsto \) locally constant. Take \( \mu_v \) rational.
- Define \( R_n(X) = \{ \phi : \mathcal{L}_X(n) \to \mathbb{R} \} \) and
  \[ \beta : R_1(X) \to R_2(X); \varphi \mapsto (\beta \varphi)(ab) = \varphi(b) - \varphi(a) \quad \forall ab \in \mathcal{L}_X(2) \]

**Theorem** If \( f \) is a coboundary, then \( \phi_n \in \beta(R_1(X_\sigma)) \) for \( n \) large enough [Host, Durand-Host-Perrin]

\[ \leadsto \phi_n(aa) = 0 \]
In short

- Take $f = 1_v - \mu_v$
- $\phi_n$ is the vector given by the sums of the values taken by $f$ on the levels of each tower. $\phi_n \in \mathbb{R}^{\mathcal{L}_\sigma(2)}$
- **Theorem** If $f$ is a coboundary of some locally constant $h$, then $\phi_n \in \beta(R_1(X_\sigma))$ for $n$ large enough

$$ (\beta \varphi)(ab) = \varphi(b) - \varphi(a) \quad \forall ab \in \mathcal{L}_X(2) \leadsto \phi_n(aa) = 0 $$
In short

- Take $f = 1_v - \mu_v$
- $\phi_n$ is the vector given by the sums of the values taken by $f$ on the levels of each tower. $\phi_n \in \mathbb{R}^{\mathcal{L}_{\sigma}(2)}$
- **Theorem** If $f$ is a coboundary of some locally constant $h$, then $\phi_n \in \beta(R_1(X_\sigma))$ for $n$ large enough

\[(\beta \varphi)(ab) = \varphi(b) - \varphi(a) \quad \forall ab \in \mathcal{L}_X(2) \sim \phi_n(aa) = 0\]

The subspace $\beta(R_1(X_\sigma))$ is easy to handle

Thue-Morse substitution

$\mathcal{L}_\sigma(2) = \{00, 01, 10, 11\}$

$$\beta(R_1(X_\sigma)) = \left\langle \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} \right\rangle$$
Symbolic discrepancy for the Thue–Morse substitution

Factors of length at least 2 have unbounded discrepancy

Proof

- $\mu_v = p_v/q_v$ with $p_v = 1$, and $q_v \in \{3 \cdot 2^{m+1}, 3 \cdot 2^m\}$ [Dekking]
- Take $f = 1_v - \mu_v$
- $\phi_n(aa) = \alpha_{aa} \left(1 - \frac{p_v}{q_v}\right) - (|\sigma^n(a)| - \alpha_{aa}) \cdot \frac{p_v}{q_v}$
- $\alpha_{aa} = \text{number of levels in the } aa \text{–tower in which all elements begin with } v$
- $\phi_n \in \beta(R_1(X)) = \langle(0, 1, -1, 0)\rangle \Rightarrow \phi_n(aa) = 0$
Symbolic discrepancy for the Thue–Morse substitution

Factors of length at least 2 have unbounded discrepancy

Proof

- \( \mu_v = p_v/q_v \) with \( p_v = 1 \), and \( q_v \in \{ 3 \cdot 2^{m+1}, 3 \cdot 2^m \} \) [Dekking]
- Take \( f = 1_v - \mu_v \)
- \( \phi_n(aa) = \alpha_{aa} \left( 1 - \frac{p_v}{q_v} \right) - (|\sigma^n(a)| - \alpha_{aa}) \cdot \frac{p_v}{q_v} \)
  \( \alpha_{aa} = \text{number of levels in the } aa\text{-tower in which all elements begin with } v \)
- \( \phi_n \in \beta(R_1(X)) = \langle (0, 1, -1, 0) \rangle \Rightarrow \phi_n(aa) = 0 \)
- \( q_v\alpha_{aa} = p_v\sigma^n(a) = p_v2^n \sim \text{Contradiction!} \)

\( \sim \) Other criteria for constant-length substitutions [B.-Cecchi]