Counting balanced words and related problems

Shigeki Akiyama (Univ. Tsukuba)

One world Numeration Seminar on 8 June 2021

Special thanks to my colleague Hiroshi Mikawa

for Large Sieve technique.
Introduction

\[ \mathcal{A} = \{0, 1\} \] and define \( \mathcal{A}^* \), \( \mathcal{A}^\mathbb{N} \) as usual. The length of \( x \in \mathcal{A}^* \) is \( |x| \). We denote \( |x|_1 \) the cardinality of 1 in \( x \in \mathcal{A}^* \). If \( y \in \mathcal{A}^* \) is a subword of \( x \in \mathcal{A}^* \cup \mathcal{A}^\mathbb{N} \), we write \( y \prec x \). A word in \( x \in \mathcal{A}^* \cup \mathcal{A}^\mathbb{N} \) is balanced if \( |u|_1 - |v|_1 | \leq 1 \) holds for any \( u, v \prec x \) with \( |u| = |v| \).

1001010 is balanced

1010001 is not balanced

An infinite word \( w \in \mathcal{A}^\mathbb{N} \) is sturmian if \( \text{Card}\{u \prec w \mid |u| = n\} = n + 1 \) for all \( n \in \mathbb{N} \).
Morse and Hedlund showed that if \( x \in \mathcal{A}^\mathbb{N} \), then followings are equivalent

- \( x \) is sturmian.
- \( x \) is aperiodic and balanced.
- \( x \) is a mechanical word of irrational slope.
A lower mechanical word \((s_n) \in \mathcal{A}^\mathbb{N}\) is defined by

\[
    s_n(\alpha, \rho) = \lfloor \alpha n + \rho \rfloor - \lfloor \alpha (n - 1) + \rho \rfloor
\]

with a given slope \(\alpha \in [0,1]\) and an intercept \(\rho \in [0,1)\). An upper mechanical word is similarly defined by replacing \(\lfloor \cdot \rfloor\) with \(\lceil \cdot \rceil\), which is denoted by \(\hat{s}_n(\alpha, \rho)\).

For every balanced word \(x = x_1 \ldots x_n\) we can find a slope \(\alpha\) and an intercept \(\rho\) such that \(x_i = s_i(\alpha, \rho) = \hat{s}_i(\alpha, \rho)\). Conversely, \((s_n(\alpha, \rho))_{n=1,\ldots,\ell}\) is balanced.
Figure 1: Balanced words
In other words, balanced words are coding of rotation.

Figure 2: Coding of rotation
Figure 3: Two interval exchange
Figure 4: Cylinder sets of rotation
Yasutomi’s partition

Let $m$ be a fixed positive integer and put $X := [0, 1) \times [0, 1]$. The map

$\psi : (\rho, \alpha) \rightarrow (s_n(1 - \alpha, \rho))_{n=1}^m$

$$X := \bigcup_{w \in B(m)} \psi^{-1}(w).$$

gives the partition:

$$Y := X \setminus \{(x, y) \mid x = ny - \ell, n \in \{1, \ldots, m\}, \ell \in \{0, 1, \ldots, n - 1\}\}.$$

The slice at $y = 1 - \alpha$ gives the cylinder partition of slope $\alpha$. Therefore $X$ is the "union" of all rotation. Berstel-Pocchiola [1] and Kamae-Takahashi [7] arrived at the same partition but this way seems to the most direct treatise.
Figure 5: Partition for $m = 4$
Lagrange spectrum appears in
\[
\sup \left\{ c > 0 \left| \left| x - \frac{p}{q} \right| < \frac{1}{cq^2} \right. \right\} \text{ has infinitely many solutions } \frac{p}{q},
\]
for any \( x \in \mathbb{R} \setminus \mathbb{Q} \), i.e.,
\[
\nu(x) = \frac{1}{\liminf_{n \to \infty} n\|nx\|} \in [\sqrt{5}, \infty].
\]
The set \( \{ \nu(x) \mid x \in \mathbb{R} \setminus \mathbb{Q} \} \) has the minimum accumulation point 3. Below 3, the discrete spectrum is described by balanced words of rational slope and intercept 0 (Christoffel words). Yasutomi [11] gave a characterization of \( x \) with \( \nu(x) = 3 \). Such words may not have a fixed slope and can be understood as a coding of 2-dim dynamics in \( X \).
Denote by \( B(n, t, u) \) the cardinality of the set of balanced words of length \( n \) whose slope \( \alpha \in [1 - t, 1] \) (Caution !) and its intercept \( \rho \in [0, u) \).

To enumerate \( B(n, t, u) \), we compute \( B(m, t, u) - B(m - 1, t, u) \), i.e., we count the number of intersections in \([0, u) \times [0, t]\) which appear by adding new segments of slope \( 1/m \), see Figure 6.

For \( t = 0.7 \) and \( u = 0.59 \), we have \((A(m, t, u))_{m=1}^{8} = (1, 2, 4, 4, 7, 8, 10, 13)\) and \( B(8, t, u) = 50 \).
Figure 6: $B(8, 0.7, 0.59) = 50$

$$B(n, t, u) = 1 + \sum_{m \leq n} A(m, t, u)$$

with

$$A(m, t, u) = \sum_{0 \leq i < j \leq m, (i,j)=1} 1,$$

where $i$ and $j$ are non negative integers.

- $(i, j) = 1$ : Möbius inversion
- $i/j \leq t$ : Easy
- $\langle mi/j \rangle < u$ : Large sieve inequality
Let \( \phi(n) \) be the Euler totient function, putting \( t = u = 1 \),

\[
A(m, 1, 1) = \sum_{0 \leq i < j \leq m} 1 = \sum_{j=1}^{m} \sum_{0 \leq i < j} 1 = \sum_{j=1}^{m} \phi(j) =: \Phi(m).
\]

\[
B(n, 1, 1) = 1 + \sum_{m=1}^{n} \Phi(m)
\]

\[
= 1 + \sum_{j=1}^{n} (n + 1 - j)\phi(j)
\]

Several different proofs are found by Lipatov, Berstel-Pocchiola, A.d.Luca-Mignosi [9, 8, 1, 3, 10].
Theorem 2.

\[ B(n, t, u) = \frac{tu}{\pi^2} n^3 + O \left( n^2 (\log n)^{15/2} \right). \]

Moreover, we have

\[ B(n, 1, 1) = \frac{n^3 + 3n^2}{\pi^2} + O \left( n^2 \exp \left( -c \left( (\log n)^{3/5} (\log \log x)^{-1/5} \right) \right) \right) \]

and

\[ B(n, t, 1) = \frac{tn^3}{\pi^2} + O(n^2). \]
Let $\mathcal{B}(n)$ be the set of balanced words of length $n$.

$$\# \{ x \in \mathcal{B}(n) \mid (\rho, \alpha) \in (a, b] \times [c, d) \} = \frac{(b - a)(d - c)}{\pi^2} n^3 + O \left( n^2 (\log n)^{15/2} \right).$$

For a Jordan measurable region $W$ in the unit square, we have Corollary 3.

$$\# \{ x \in \mathcal{B}(n) \mid (\rho, \alpha) \in W \} = \frac{\text{Area}(W)}{\pi^2} n^3 + O \left( n^2 (\log n)^{15/2} \right)$$

where \text{Area} is the 2-dimensional Lebesgue measure.
Farey series \((f_m(i))\) of order \(m\) is the finite increasing sequence composed of irreducible fractions in \([0, 1)\) whose denominators are not larger than \(m\):

\[0 = f_m(1) < f_m(2) < \cdots < f_m(\Phi(m)) < 1.\]

For \(m = 6\), we have

\[
\begin{align*}
&0 < \frac{1}{6} < \frac{1}{5} < \frac{1}{4} < \frac{1}{3} < \frac{2}{5} < \frac{1}{2} < \frac{3}{5} < \frac{2}{3} < \frac{3}{4} < \frac{4}{5} < \frac{5}{6} < 1
\end{align*}
\]

Clearly \(A(m, t, 1) = \max \{j \mid f_m(j) \leq t\}\). For example

\[A \left(6, \frac{1}{2}, 1 \right) = 7.\]
The function $A(m, t, u)$ counts special Farey fractions. An easy estimate

$$A(m, t, 1) - t\Phi(m) = O(m)$$

gives uniform distribution property. Farey series is expected to be "highly" uniform in $[0, 1]$. J. Franel showed

$$\text{Riemann Hypothesis} \iff \int_0^1 (A(m, t, 1) - t\Phi(m))^2 \, dt = O\left(m^{1+\varepsilon}\right).$$

We add another one directly related to the number of balanced words:

**Corollary 4.**

$$\text{Riemann Hypothesis} \iff B(n, 1, 1) = \frac{n^3 + 3n^2}{\pi^2} + O\left(n^{3/2+\varepsilon}\right). \quad (1)$$
Counting technique in number theory

An arithmetic functions is the map $\mathbb{N} \to \mathbb{C}$. With binary operations

$$(f + g)(n) = f(n) + g(n),$$

and

$$(f * g)(n) = \sum_{d|n} f(d)g \left( \frac{n}{d} \right), \quad \text{( Dirichlet convolution )}$$

they form a commutative ring $\mathcal{R}$ with the multiplicative identity $e$:

$$e(n) = \begin{cases} 
1 & n = 1 \\
0 & n > 1
\end{cases}$$
The invertible elements are \( \mathcal{R}^* = \{ f \in \mathcal{R} \mid f(1) \neq 0 \} \). Denote by \( \mathbf{1} \) the function \( \mathbf{1}(n) = 1 \). Then we have

\[
\mathbf{1} \ast \mu = e
\]

where

\[
\mu(n) = \begin{cases} 
1 & n = 1 \\
(-1)^k & n = p_1 \ldots p_k \ (p_i : \text{distinct primes}) \\
0 & n \text{ is not square free}
\end{cases}
\]

Möbius inversion formula is well-known.

\[
f = g \ast \mathbf{1} \iff f \ast \mu = g
\]

i.e.

\[
f(n) = \sum_{d | n} g(d) \iff g(n) = \sum_{d | n} f(d) \mu \left( \frac{n}{d} \right)
\]
We often study its generating Dirichlet series

\[ D(f, s) := \sum_{n=1}^{\infty} \frac{f(n)}{n^s}. \]

Then we have formally

\[ D(f \ast g, s) = D(f, s)D(g, s). \]

Many important arithmetic functions are connected with the Riemann zeta function.

\[ D(1, s) = \zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_p \frac{1}{1 - \frac{1}{p^s}}. \]
For e.g.

\[ D(\mu, s) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \frac{1}{\zeta(s)}, \quad D(\phi, s) = \sum_{n=1}^{\infty} \frac{\phi(n)}{n^s} = \frac{\zeta(s - 1)}{\zeta(s)} \]

We use properties of Riemann zeta function as an analytic function in \( \mathbb{C} \) to deduce asymptotic formula of summatory function, by Perron's formula:

\[
\sum_{n<x} f(n) = \lim_{T \to \infty} \frac{1}{2\pi \sqrt{-1}} \int_{c-T \sqrt{-1}}^{c+T \sqrt{-1}} \left( \sum_{n=1}^{\infty} \frac{f(n)}{n^s} \right) \frac{x^s}{s} ds
\]

or its weighted variants. The most famous one is the Prime number theorem:

\[
\sum_{p<x} \frac{1}{\log x} + O(x \exp(-c \sqrt{\log x})).
\]
Case \( t = u = 1 \).

\[
B(n, 1, 1) = 1 + \sum_{j=1}^{n} (n + 1 - j) \phi(j).
\]

We use

\[
\sum_{j \leq x} (x - j) \phi(j) = \lim_{T \to \infty} \frac{1}{2\pi \sqrt{-1}} \int_{a - \sqrt{-1}T}^{a + \sqrt{-1}T} \frac{\zeta(s - 1)x^{s+k}}{\zeta(s)s(s+1)} ds
\]

for \( a > 2 \). For Corollary 4, if Riemann Hypothesis is valid,

\[
\lim_{T \to \infty} \int_{a_0 - \sqrt{-1}T}^{a_0 + \sqrt{-1}T} \left| \frac{\zeta(s - 1)}{\zeta(s)s(s+1)} \right| ds < \infty,
\]
with \( a_0 = 1/2 + \varepsilon \), we get the estimate (1). Conversely inverse Mellin transformation shows

\[
\int_{1}^{\infty} \left( \sum_{n \leq x} (x - n) \phi(n) - \frac{x^3}{\pi^2} \right) x^{-s-2} \, dx = \frac{\zeta(s-1)}{\zeta(s) s(s+1)} - \frac{1}{\pi^2(s-2)}
\]

for \( \sigma > 2 \). If (1) is valid, then the parenthesis in the integrand is \( O(x^{3/2+\varepsilon}) \). This gives the holomorphic continuation of the right side to \( \sigma > 1/2 + \varepsilon \), which finishes the proof. Without RH, we use the current best "zero free" region by shifting the path.
\[ \delta = 0 \quad \delta = \frac{1}{2} \quad \delta = 1 \quad \delta = 2 \]

\[ \zeta(s-1)/(\zeta(s)s(s+1)) \]

Figure 7: Shifting the integration path
to obtain

$$B(n, 1, 1) = \frac{n^3 + 3n^2}{\pi^2} + O \left( n^2 \exp \left( -c \left( (\log n)^{3/5} (\log \log x)^{-1/5} \right) \right) \right).$$
Case $u = 1$.

$$B(n, t, 1) = 1 + \sum_{m \leq n} A(m, t, 1), \quad A(m, t, u) = \sum_{\substack{i < j \leq m, \ (i, j) = 1 \\ i/j \leq t}} 1.$$ 

The key equation is

$$B(n, t, 1) = 1 - t + n + tB(n, 1, 1) - \sum_{kb \leq m \leq n} \mu(k)\langle bt\rangle$$

for $t < 1$. 
Our problem is the summatory function:

\[
\sum_{kb \leq m} \mu(k) \langle bt \rangle = \sum_{\ell=1}^{m} \sum_{k|\ell} \mu(k) \left\langle \frac{\ell}{k} t \right\rangle
\]

Therefore the analytic property of

\[
Z_t(s) / \zeta(s)
\]

plays a key role where

\[
Z_t(s) := \sum_{b=1}^{\infty} \frac{\langle bt \rangle - \frac{1}{2}}{b^s}.
\]

is the Hecke’s Dirichlet series.
Fujii [4] studied Hecke’s Dirichlet series [5]: The analytic property of $Z_t(s)$ heavily depends on the Diophantine approximation property of $t$ by rationals. [4, Theorem 1 and 2] imply

$$B(n, t, 1) = \frac{t(n^3 + 3n^2)}{\pi^2} + O\left(n^2 \exp\left(-c (\log n \cdot \log \log n)^{1/3}\right)\right)$$

for almost all $t$, including all algebraic numbers.

However for general $t$, the error term is pretty big for now.

$$B(n, t, 1) = \frac{tn^3}{\pi^2} + O(n^2).$$
General case.

\[
B(n, t, u) = 1 + \sum_{m \leq n} A(m, t, u), \quad A(m, t, u) = \sum_{\substack{i < j \leq m, \ (i,j) = 1 \\ i/j \leq t, \ \langle mi/j \rangle < u}} 1.
\]

- Fourier expansion of a smooth approximation of \( \chi[0,u] \mod \mathbb{Z} \).
- Remove other constraints by Fourier technique.
- Apply large sieve inequality to obtain cancellation.
Fourier expansion

Choose $\chi[0,u] \pmod{\mathbb{Z}} \approx \sum_{h \in \mathbb{Z}} a_h \exp \left( \frac{2\pi m h \sqrt{-1}}{j} \right)$ to obtain

$$
\sum_{i<j \leq m, \ (i,j)=1} 1 \approx \sum_{i<j \leq m, \ (i,j)=1} \sum_{h \in \mathbb{Z}} a_h \exp \left( \frac{2\pi m h \sqrt{-1}}{j} \right)
$$

$$
\approx \sum_{i<j \leq m, \ (i,j)=1} \sum_{\substack{h=-H \ \text{to} \ h=H}} a_h \exp \left( \frac{2\pi m h \sqrt{-1}}{j} \right)
$$

Constant term $h = 0$ gives the main term.
Remove other constraints

Use Möbius inversion to delete \((i, j) = 1\). Remove \(i/j \leq t\) by

\[
\frac{2}{\pi} \int_0^T \sin \frac{x}{x} dx = 1 + O \left( \min \left( 1, \frac{1}{T} \right) \right),
\]

The target becomes

\[
\sum_{i<j \leq m} \sum_{h=1}^H a_h \exp \left( 2\pi mih \sqrt{-1} \frac{j}{j} \right)
\]
Apply large sieve inequality

Large sieve stands for a certain variants of Bessel inequality. Letting \( S(\alpha) = \sum_{n=1}^{N} a_n e(n\alpha) \), and \( \alpha_1, \ldots, \alpha_R \in \mathbb{T} \) with \( d(\alpha_i, \alpha_j) \geq \delta \) for \( i \neq j \), we have

\[
\sum_{r=1}^{R} |S(\alpha_r)|^2 \leq \left( N + \frac{1}{\delta} - 1 \right) \sum_{n=1}^{N} |a_n|^2
\]

is a well-known one by Halberstam-Davenport-Selberg.

There are many many variants but I had a Hiroshi Mikawa’s advice.
We shall use a large sieve inequality [6, Theorem 7.2],[2, Lemma 2.4]:

**Lemma 5.** For any real numbers $x_m, y_m$ with $|x_m| \leq X$ and $|y_m| \leq Y$ and $\alpha_m, \beta_m \in \mathbb{C}$, we have

\[
\left| \sum_{m} \sum_{n} \alpha_m \beta_n \exp(2\pi x_m y_n \sqrt{-1}) \right| \leq 5\sqrt{1 + XY} \left( \sum_{|x_i - x_j| < 1/Y} |\alpha_i \alpha_j| \sum_{|y_i - y_j| < 1/X} |\beta_i \beta_j| \right)^{1/2}.
\]

We obtain

\[ B(n, t, u) = \frac{tu}{\pi^2 n^3} + O \left( n^2 (\log n)^{15/2} \right) \]
by the choice

\[ x_n = m i, \quad y_m = \frac{h}{j}. \]

References


