Abelian Ramsey Length and Asymptotic Lower Bounds*

Vincent Jugé†

This technical note aims at evaluating an asymptotic lower bound on abelian Ramsey lengths obtained by Tao in [1]. We first provide the minimal amount of background necessary to define abelian Ramsey lengths, and indicate the lower bound of Tao. We then focus on evaluating this lower bound.

1 Introduction

Let \( \mathcal{A} \) and \( \mathcal{V} \) be two alphabets. A word on \( \mathcal{A} \) is a finite sequence \( a = a_1 a_2 \cdots a_k \) of elements of \( \mathcal{A} \). The elements \( a_i \) are called the letters of the word \( a \), and the integer \( k \) is the length of \( a \). For all elements \( \alpha \in \mathcal{A} \), we denote by \( |a|_\alpha \) the cardinality of the set \( \{ i : a_i = \alpha \} \), i.e. number of occurrences of the letter \( \alpha \) in the word \( a \). We also denote by \( \mathcal{A}^* \) the set of all words on \( \mathcal{A} \). The words \( a_i a_{i+1} \cdots a_j \) with \( 1 \leq i \leq j \leq k \), as well as the empty word, are called factors of \( a \).

Consider now a word \( a = a_1 a_2 \cdots a_k \) in \( \mathcal{A}^* \) and a word \( p = p_1 p_2 \cdots p_l \) in \( \mathcal{V}^* \). We say that \( a \) contains \( p \) in the abelian sense if there exist non-empty words \( \pi_1, \pi_2, \ldots, \pi_\ell \in \mathcal{A}^* \) such that the concatenated word \( \pi_1 \pi_2 \cdots \pi_\ell \) is a factor of \( a \), and such that, for all integers \( i, j \) and all letters \( \alpha \in \mathcal{A} \), if \( p_i = p_j \), then \( |\pi_i|_\alpha = |\pi_j|_\alpha \). For instance, the word "programmable" contains the word "aab" in the abelian sense, as can be seen by considering the words "\( a = am \), \( \pi_2 = ma \) and \( \pi_3 = ble \)."

From this point on, we consider the infinite alphabet \( \mathcal{V} = \{ v_i : i \in \mathbb{N} \} \), where \( \mathbb{N} \) is the set of positive integers, and we define the Zimin patterns \( Z_i \) inductively by \( Z_1 = v_1 \) and \( Z_{i+1} = Z_i v_{i+1} Z_i \). It turns out that, for all integers \( i, m \geq 1 \) and all alphabets \( \mathcal{A} \) of cardinality \( m \), there exists an integer \( L_{ab}(m, Z_i) \) such that all words \( a \in \mathcal{A}^* \) with length at least \( L_{ab}(m, Z_i) \) contain the word \( Z_i \) in the abelian sense.

For all integers \( m \geq 4 \), Tao proves in [1] that \( L_{ab}(m, Z_i) \geq (1 + \varepsilon_m(i))\sqrt{K(m, i)} \) for all \( i \geq 1 \), where \( \varepsilon_m \) is a function such that \( \lim_{+\infty} \varepsilon_m = 0 \) and \( K(m, i) \) is defined as

\[
K(m, i) = 2 \prod_{j=1}^{i-1} S(m, 2^j)^{-1},
\]

where

\[
S(m, k) = \sum_{\mathclap{\ell=1}}^{\infty} T(m, k, \ell) \quad \text{and} \quad T(m, k, \ell) = \frac{1}{m^{\ell}} \sum_{\mathclap{i_1 + \cdots + i_m = \ell}} \left( i_1 \cdots i_m \right)^k.
\]

Yet, in order to obtain actual lower bounds on \( L_{ab}(m, Z_i) \), it remains to evaluate the asymptotical behavior of \( K(m, i) \). We evaluate \( K(m, i) \) up to a multiplicative constant that does not depend on \( m \) or \( i \). More precisely, we prove the following inequalities, which hold for all \( m \geq 4 \) and \( i \geq 1 \):

\[
2 \frac{m^{2i}}{m^{i+1}} \geq K(m, i) \geq \frac{1}{21} \frac{m^{2i}}{m^{i+1}}.
\]

---

*This work is supported by ERC EQualIS (308087).
†LSV, CNRS & ENS Cachan, Univ. Paris-Saclay, France
2 Auxiliary inequalities

Before evaluating the lower bound $K(m, i)$, we prove a series of six inequalities that we will use subsequently. We first study the function $f_x : y \mapsto y \ln(1 + x/y)$ for $x, y > 0$. An asymptotic evaluation proves that \( \lim_{y \to \infty} f_x = x \). Furthermore, we compute that $f'_x(y) = -\frac{x^2}{y(x+y)^2} < 0$ for $y > 0$. It follows that $x > f_x(y)$ or, equivalently, that

\[
(1 + x/y)^y < e^x \quad \text{for all } x, y > 0. \tag{1}
\]

We perform a similar study with the function $g : y \mapsto (y + 1/2) \ln(1 + 1/y)$ for $y > 0$. We find that $\lim_{y \to \infty} g = 1$ and that $g''(y) = \frac{1}{y^2(1+y^2)} > 0$ for $y > 0$. It follows that $g(y) > 1$ or, equivalently, that

\[
(1 + 1/y)^{y+1/2} > e \quad \text{for all } y > 0. \tag{2}
\]

Again, we consider the function $h : y \mapsto 3 \ln(y) + \ln(2) - (y - 1) \ln(2\pi) + 7$ for $y > 0$, as well as the real constant $\lambda = \frac{\ln(2)}{\sqrt{\pi}} < \frac{1}{e}$. We find that $h'(y) = \frac{3}{y} - \ln(2\pi) < 0$ when $y \geq 4$ and that $\exp(h(4)) = \frac{16}{\pi^3} = \lambda^2$, and it follows that

\[
2y^3 \leq \lambda^2(2\pi)^{y-1} \quad \text{for all } y \geq 4. \tag{3}
\]

Similarly consider the function $\overline{h} : y \mapsto 5 \ln(y) + \ln(2) - (y - 1) \ln(2\pi) + 7$ for $y > 0$. We find that $\overline{h}(y) = \frac{5}{y} - \ln(2\pi) < 0$ when $y \geq 7$ and that $\exp(\overline{h}(7)) = \frac{5^3}{\pi^3} < 1$, and it follows that

\[
2y^5 \leq (2\pi)^{y-1} \quad \text{for all } y \geq 7. \tag{4}
\]

Then, we set $Z(x) = \sqrt{2\pi x^{x+1/2}e^{-x}}$ for all $x \geq 0$. We prove below that

\[
\frac{(a+b)!}{ab!} \leq \frac{Z(a)Z(b)}{Z(a+b)} \quad \text{for all integers } a, b \geq 1. \tag{5}
\]

We study the functions $F : (a, b) \mapsto \frac{(a+b)!Z(a)Z(b)}{Z(a+b)a!b!}$ and $G : (a, b) \mapsto \frac{F(a+1, b)}{F(a, b)}$. We compute that

\[
G(a, b) = \left( \frac{a+b}{a+b+1} \right)^{a+b+1/2} \left( \frac{a+1}{a} \right)^{a+3/2} \geq \overline{G}(a, b)^{2a+b+2}, \quad \text{where}
\]

\[
\overline{G}(a, b) = \frac{2a+b+2}{(a+b+1/2)^{a+1} + (a+3/2)^{a+1}} \quad \text{(by geometric-harmonic inequality)}
\]

\[
= 1 + \frac{b-1}{2a+b+1}(a+1) + (a+3)a(a+b).
\]

and since $b \geq 1$, it follows that $G(a, b) \geq \overline{G}(a, b)^{2a+b+2} \geq 1$, i.e. that $F(a, b) \leq F(a+1, b)$. Since $F(a, b) = F(b, a)$ for all $a, b \geq 1$, we derive immediately that $F(a, b) \leq F(a+b+1, b) \leq F(a+1, b+1)$ for all integers $a, b \geq 1$. Moreover, Stirling’s approximation formula states that $a! \sim Z(a)$ when $a \to +\infty$. This proves that $\lim_{a, b \to +\infty} F(a, b) = 1$, and it follows that $F(a, b) \leq 1$ for all $a, b \geq 1$, which is indeed equivalent to the inequality (5).

As a corollary, observe that, for all integers $i_1, \ldots, i_m \geq 1$ and using inequality (5), we also have

\[
\frac{(i_1 + \ldots + i_m)!}{i_1! \ldots i_m!} = \prod_{j=2}^{m} \frac{(i_1 + \ldots + i_{j-1})!i_j!}{Z(i_1 + \ldots + i_{j-1})Z(i_j)} \leq \prod_{j=2}^{m} \frac{Z(i_1 + \ldots + i_j)}{Z(i_1 + \ldots + i_{j-1})Z(i_j)},
\]

from which follows our last auxiliary inequality:

\[
\frac{(i_1 + \ldots + i_m)!}{i_1! \ldots i_m!} \leq \frac{Z(i_1 + \ldots + i_m)}{Z(i_1) \ldots Z(i_m)} \quad \text{for all integers } i_1, \ldots, i_m \geq 1. \tag{6}
\]
3 Evaluating $K(m, i)$

We first evaluate $T(m, k, \ell)$ when $\ell = 1$. Here, instead of considering a tuple of non-negative integers $(i_1, \ldots, i_m)$ that sum up to $\ell$, we might directly consider the unique integer $j \in \{1, \ldots, m\}$ such that $i_j = 1$. Moreover, for each tuple $(i_1, \ldots, i_m)$, the multinomial coefficient $(i_1 \ldots i_m)$ is equal to 1. It follows that $T(m, k, 1) = m^{-k} \sum_{i=1}^{m} 1 = m^{1-k}$, from which we derive the inequalities $S(m, k) \geq T(m, k, 1) = m^{1-k}$ and

$$K(m, i) \leq 2 \prod_{j=1}^{i-1} m^{2^j-1} = 2 \frac{m^2}{m^{i+1}}.$$

Then, we investigate lower bounds of $K(m, i)$, i.e. upper bounds of $T(m, k + 1, \ell)$ and of $S(m, k + 1)$ when $m \geq 4$ and $k \geq 1$. Consider some integer $\ell \geq 1$, and let us write $\ell = am + b$, with $a \geq 0$ and $1 \leq b \leq m$. In addition, let us set

$$V(m, a, b) = \max_{i_1 + \ldots + i_m = \ell} \binom{\ell}{i_1 \ldots i_m} \text{ and } U(m, a, b) = \frac{1}{m^a} V(m, a, b).$$

We first observe that

$$T(m, k + 1, \ell) = \frac{1}{m^{k+1}} \sum_{i_1 + \ldots + i_m = \ell} \binom{\ell}{i_1 \ldots i_m}^{k+1} \leq \frac{1}{m^{k+1}} \sum_{i_1 + \ldots + i_m = \ell} \binom{\ell}{i_1 \ldots i_m} V(m, a, b)^k \leq U(m, a, b)^k T(m, 1, \ell) \leq U(m, a, b)^k. \quad \text{(by Newton multinomial identity)}$$

Since the inequality $(x + 1)! (y - 1)! \geq x! y!$ holds for all integers $x \geq y$, it also follows that

$$U(m, a, b) = \frac{(am + b)!}{m^{am+b}(a + 1)!a!b!m^b}.$$

We compute immediately that $U(m, 0, b) = m^{-b}b! = m^{-1}$ if $b = 1$, and that $U(m, 0, b) \leq 2m^{-2}$ if $2 \leq b \leq m$. When $a \geq 1$, we further compute that

$$U(m, a, b) \leq \frac{Z(am + b)}{m^{am+b}Z(a + 1)!Z(a)^{m-b}} \leq \frac{\sqrt{am + b}}{(2\pi)^{m-1}/2m^m/2} \frac{(1 + b/am)^{am}(1 + b/am)^b}{(1 + 1/a)^{am}(1 + 1/a)^b} \leq \frac{\sqrt{2m}}{(2\pi)^{(m-1)/2am^m/2}} \frac{(1 + b/am)^m}{(1 + 1/a)^{am}(1 + 1/a)^b} \leq \frac{\sqrt{2m}}{(2\pi)^{(m-1)/2}} \frac{e^b}{(1 + 1/a)^{am}(1 + 1/a)^b} \leq \frac{\sqrt{2m}}{(2\pi)^{(m-1)/2}} \left( \frac{e}{(1 + 1/a)^{am}(1 + 1/a)^b} \right)^b \leq \frac{\sqrt{2m}}{(2\pi)^{(m-1)/2}}. \quad \text{(using inequality (1))}$$

$$U(m, a, b) \leq \frac{\sqrt{2m}}{(2\pi)^{(m-1)/2}}. \quad \text{(using inequality (2))}$$
Consequently, since \( k \geq 1 \), we find that

\[
S(m, k + 1) \leq \sum_{a=0}^{\infty} \sum_{b=1}^{m} U(m, a, b)^k = U(m, 0, 1)^k + \sum_{b=2}^{m} U(m, b)^k + \sum_{a=1}^{\infty} \sum_{b=1}^{m} U(m, a, b)^k
\]

\[
\leq \frac{1}{m^k} + \frac{2(m-1)}{m^{2k}} + m \sum_{a=1}^{\infty} \left( \frac{\sqrt{2m}}{(2\pi)(m-1/2)} \right)^k
\]

\[
\leq \frac{1}{m^k} + \frac{2m}{m^{2k}} + m \frac{(2m)^{k/2}}{(2\pi)k^{k/2}} \zeta(k(m-1)/2).
\]

If we set

\[
P(m, k) = 1 + 2m^{1-k} + m^{k+1} \frac{(2m)^{k/2}}{(2\pi)k^{k/2}} \zeta(k(m-1)/2),
\]

then it follows that \( S(m, k + 1) \leq \frac{1}{m} P(m, k) \) and therefore that

\[
K(m, i) \geq 2 \prod_{j=1}^{i-1} \frac{m^{2j-1}}{P(m, 2^j - 1)} \geq \frac{2}{P_\infty(m)} \frac{m^{2i}}{m^{i+1}},
\]

where \( P_\infty(m) \) is the infinite product \( \prod_{j=1}^{\infty} P(m, 2^j - 1) \). It remains to prove that \( P_\infty(m) \leq 42 \).

We first assume that \( 7 \leq m \). For \( k \geq 1 \), we compute that

\[
P(m, k) = 1 + 2m^{1-k} + m^{k+1} \frac{(2m)^{k/2}}{(2\pi)k^{k/2}} \zeta(k(m-1)/2)
\]

\[
= 1 + 2m^{1-k} + m^{k-1} \left( \frac{2m^5}{(2\pi)^3 m^{-1}} \right)^{k/2} \zeta(k(m-1)/2)
\]

\[
\leq 1 + 2m^{1-k} + m^{1-k} \zeta(k(m-1)/2)
\]

(\text{using inequality (4)})

\[
\leq 1 + 4m^{1-k}
\]

(\text{since } \zeta(k(m-1)/2) \leq \zeta(3) \leq 2)

\[
\leq \exp(4m^{1-k})
\]

(\text{since } 1 + x \leq \exp(x) \text{ for all } x \in \mathbb{R})

from which we deduce that \( P_\infty(m, 1) \leq 5 \), whence

\[
\ln(P_\infty(m)) = \sum_{j=1}^{\infty} \ln(P(m, 2^j - 1)) \leq \ln(P(m, 1)) + 4 \sum_{j=2}^{\infty} m^{2-2^j} \leq \ln(5) + 4 \sum_{j=0}^{\infty} m^{-2^j}
\]

\[
\leq \ln(5) + \frac{4}{m(m-1)} \leq \ln(5) + \frac{2}{21} \leq \ln(42).
\]

(since \( 7 \leq m \))

Then, we assume that \( 4 \leq m \leq 6 \). Again, for \( k \geq 1 \), we compute that

\[
P(m, k) = 1 + 2m^{1-k} + m^{k+1} \frac{(2m)^{k/2}}{(2\pi)k^{k/2}} \zeta(k(m-1)/2)
\]

\[
= 1 + 2m^{1-k} + m \left( \frac{2m^3}{(2\pi)^3 m^{-1}} \right)^{k/2} \zeta(k(m-1)/2)
\]

(\text{using inequality (3)})

\[
\leq 1 + 2m^{1-k} + m \zeta(k(m-1)/2) \lambda^k
\]

\[
\leq 1 + 2m^{1-k} + 3m \lambda^k
\]

(\text{\( \zeta(k(m-1)/2) \leq \zeta(3)/2 \leq 3 \)})

\[
\leq 1 + 5m \lambda^k
\]

(\text{since } m^{-1} \leq \frac{1}{\lambda} \leq \lambda)

\[
\leq \exp(5m \lambda^k)
\]

(\text{since } 1 + x \leq \exp(x) \text{ for all } x \in \mathbb{R})
Furthermore, explicit computations in each of the cases $m = 4$, $m = 5$ and $m = 6$ indicate that $\prod_{j=1}^{4} P(m, 2^j - 1) \leq 41$. Hence, we conclude that

$$\ln(P_\infty(m)) \leq \ln(41) + \sum_{j=5}^{\infty} \ln(P(m, 2^j - 1))$$

$$\leq \ln(41) + 5m \sum_{j=5}^{\infty} \lambda^{2^j - 1} \leq \ln(41) + 5m \sum_{j=0}^{\infty} \lambda^{31+j}$$

$$\leq \ln(41) + \frac{5m \lambda^{31}}{1 - \lambda} \leq \ln(41) + \frac{30 \times 3^{31}}{4^{30}} \leq \ln(42). \quad \text{(since } m \leq 6 \text{ and } \lambda < \frac{4}{3})$$

References

[1] J. Tao. Pattern occurrence statistics and applications to the Ramsey theory of unavoidable patterns. *ArXiv e-prints*, June 2014.