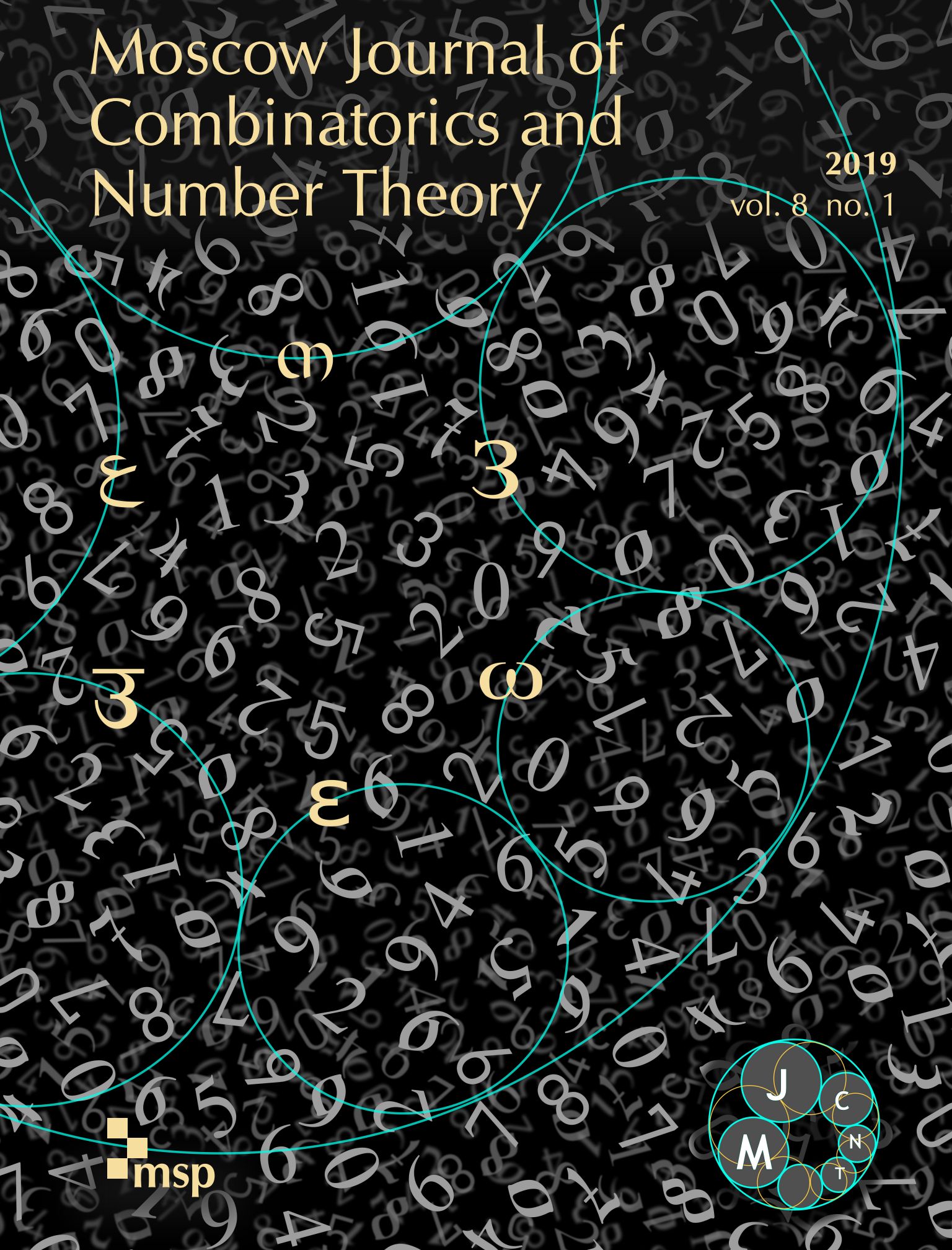
Moscow Journal of Combinatorics and Number Theory

2019
vol. 8 no. 1



msp

Moscow Journal of Combinatorics and Number Theory

msp.org/moscow

EDITORS-IN-CHIEF

Nikolay Moshchevitin	Lomonosov Moscow State University (Russia) moshchevitin@gmail.com
Andrei Raigorodskii	Moscow Institute of Physics and Technology (Russia) mraigor@yandex.ru

EDITORIAL BOARD

Yann Bugeaud	Université de Strasbourg (France)
Vladimir Dolnikov	Moscow Institute of Physics and Technology (Russia)
Nikolay Dolbilin	Steklov Mathematical Institute (Russia)
Oleg German	Moscow Lomonosov State University (Russia)
Grigory Kabatiansky	Russian Academy of Sciences (Russia)
Roman Karasev	Moscow Institute of Physics and Technology (Russia)
Gyula O. H. Katona	Hungarian Academy of Sciences (Hungary)
Alex V. Kontorovich	Rutgers University (United States)
Maxim Korolev	Steklov Mathematical Institute (Russia)
Christian Krattenthaler	Universität Wien (Austria)
Antanas Laurinčikas	Vilnius University (Lithuania)
Vsevolod Lev	University of Haifa at Oranim (Israel)
János Pach	EPFL Lausanne(Switzerland) and Rényi Institute (Hungary)
Rom Pinchasi	Israel Institute of Technology – Technion (Israel)
Alexander Razborov	Institut de Mathématiques de Luminy (France)
Joël Rivat	Université d'Aix-Marseille (France)
Tanguy Rivoal	Institut Fourier, CNRS (France)
Damien Roy	University of Ottawa (Canada)
Vladislav Salikhov	Bryansk State Technical University (Russia)
Tom Sanders	University of Oxford (United Kingdom)
Alexander A. Sapozhenko	Lomonosov Moscow State University (Russia)
Ilya D. Shkredov	Steklov Mathematical Institute (Russia)
József Solymosi	University of British Columbia (Canada)
Benjamin Sudakov	University of California, Los Angeles (United States)
Jörg Thuswaldner	University of Leoben (Austria)
Kai-Man Tsang	Hong Kong University (China)
Maryna Viazovska	EPFL Lausanne (Switzerland)

PRODUCTION

Silvio Levy	(Scientific Editor) production@msp.org

See inside back cover or msp.org/moscow for submission instructions.

The subscription price for 2019 is US $310/year for the electronic version, and $365/year (+$20, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Moscow Journal of Combinatorics and Number Theory (ISSN 2220-5438 electronic, 2220-5438 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

MJCNT peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

© 2019 Mathematical Sciences Publishers

msp

Moscow Journal of Combinatorics and Number Theory
Vol. 8, No. 1, 2019
dx.doi.org/10.2140/moscow.2019.8.1

To the reader

Dear Reader:

The *Moscow Journal of Combinatorics and Number Theory* was founded in 2010 by the Moscow Institute of Physics and Technology, and in 2018 it started being published by MSP (Mathematical Sciences Publishers), a nonprofit scientific publisher based in Berkeley, California.

Our journal publishes original, high-quality research articles from a broad range of interests within combinatorics, number theory and allied areas. Since 2011 we have published over 100 papers. Among our authors are such mathematicians as Noga Alon, Antal Balog, Jean-Pierre Demailly, Dominic Foata, Peter Frankl, Aleksandar Ivić, Sergei Konyagin, Yuri Nesterenko, János Pach, Yakov Sinai, Andrzej Schinzel, Wolfgang Schmidt, Carlo Viola, Michel Waldschmidt, and many others.

This issue 1 of volume 8 is the first issue to appear under MSP's aegis. It contains selected papers presented by the participants of the **Vilnius Conference in Combinatorics and Number Theory**, which was organized with support from our journal and took place at the Department of Mathematics and Informatics of the University of Vilnius, Lithuania, 16–22 July 2017.

Previous conferences connected to our journal were held in Russia (**Diophantine analysis**, Astrakhan, 30 July to 3 August 2012), Lithuania (**Palanga Conference in Combinatorics and Number Theory**, 1–7 September 2013), again Russia (**Moscow Workshop in Combinatorics and Number Theory**, January 27 to 2 February 2014), and Denmark (**Diophantine Approximation and Related Topics**, Aarhus, 13–17 July 2015).

A collection of papers from the Astrakhan conference appeared in issue 3–4 of volume 3 (2013), and papers related to the Aarhus conference in issue 2–3 of volume 6 (2016).

We hope that you will enjoy this issue and support the journal both with your submissions and by recommending a subscription to your institutional library!

NIKOLAY MOSHCHEVITIN:

moshchevitin@gmail.com
Lomonosov Moscow State University Moscow, Russia

ANDREI RAIGORODSKII:

mraigor@yandex.ru
Moscow institute of Physics and Technology, and Lomonosov Moscow State University Moscow, Russia

msp

Moscow Journal of Combinatorics and Number Theory
Vol. 8, No. 1, 2019
dx.doi.org/10.2140/moscow.2019.8.3

Sets of inhomogeneous linear forms can be not isotropically winning

Natalia Dyakova

We give an example of irrational vector $\boldsymbol{\theta} \in \mathbb{R}^2$ such that the set

$$\mathrm{Bad}_{\boldsymbol{\theta}} := \big\{(\eta_1, \eta_2) : \inf_{x\in\mathbb{N}} x^{1/2} \max_{i=1,2} \|x\theta_i - \eta_i\| > 0\big\}$$

is not absolutely winning with respect to McMullen's game.

1. Introduction

We consider a problem related to inhomogeneous Diophantine approximation. Given $\boldsymbol{\theta} = (\theta_1, \theta_2) \in \mathbb{R}^2$ we study the set of pairs $(\eta_1, \eta_2) \in \mathbb{R}^2$ such that the system of two linear forms

$$\|x\theta_1 - \eta_1\|, \quad \|x\theta_2 - \eta_2\|,$$

where $\|\cdot\|$ stands for the distance to the nearest integer, is badly approximable. We prove a statement complementary to our recent result from [Bengoechea et al. 2017]. We construct $\boldsymbol{\theta}$ such that the set

$$\mathrm{Bad}_{\boldsymbol{\theta}} := \big\{(\eta_1, \eta_2) : \inf_{x\in\mathbb{N}} x^{1/2} \max_{i=1,2} \|x\theta_i - \eta_i\| > 0\big\}$$

is not isotropically winning.

Our paper is organized as follows. In Section 2 we discuss different games appearing in Diophantine problems. In Section 3 we give a brief survey on inhomogeneous badly approximable systems of linear forms and formulate our main result, Theorem 3.1. Sections 4 and 5 are devoted to some auxiliary observations. In Sections 6, 7, and 8 we give a proof for Theorem 3.1.

2. Schmidt's game and its generalizations

The following game was introduced by Schmidt [1966; 1969; 1980]. Let $0 < \alpha, \beta < 1$. Suppose that two players A and B choose in turn a nested sequence of closed balls:

$$B_1 \supset A_1 \supset B_2 \supset A_2 \supset \cdots$$

with the property that the diameters $|A_i|, |B_i|$ of the balls A_i, B_i satisfy

$$|A_i| = \alpha|B_i|, \quad |B_{i+1}| = \beta|A_i| \qquad \text{for all } i = 1, 2, 3, \ldots,$$

The author is supported by RFBR Grant No. 18-01-00886a.
MSC2010: 11J13.
Keywords: inhomogeneous diophantine approximation, winning sets.

for fixed $0 < \alpha, \beta < 1$. A set $E \subset \mathbb{R}^n$ is called (α, β)-*winning* if player A has a strategy which guarantees that intersection $\bigcap A_i$ meets E regardless of the way B chooses to play. A set $E \supset \mathbb{R}^n$ is called an α-*winning* set if it is (α, β)-winning for all $0 < \beta < 1$.

There are different modifications of Schmidt's game: the strong game and absolute game introduced in [McMullen 2010], the hyperplane absolute game introduced in [Kleinbock and Weiss 2010], the potential game considered in [Fishman et al. 2013], and some others. In [Bengoechea et al. 2017], we introduced isotropically winning sets. Let us describe here some of these generalizations in more detail.

The definition of an absolutely winning set was given in [McMullen 2010]. Consider the following game. Suppose A and B choose in turn a sequence of balls A_i and B_i such that the sets

$$B_1 \supset (B_1 \backslash A_1) \supset B_2 \supset (B_2 \backslash A_2) \supset B_3 \supset \cdots$$

are nested. For fixed $0 < \beta < \frac{1}{3}$ we suppose

$$|B_{i+1}| \geq \beta |B_i|, \quad |A_i| \leq \beta |B_i|.$$

We say E is an *absolute winning* set if for all $\beta \in \left(0, \frac{1}{3}\right)$, player A has a strategy which guarantees that $\cap B_i$ meets E regardless of how B chooses to play. Mcmullen proved that an absolute winning set is α-winning for all $\alpha < \frac{1}{2}$. Several examples of absolute winning sets were exhibited by McMullen [2010]. In particular, a set of badly approximable numbers in $\mathbb{R}$ is absolutely winning. However the set of simultaneously badly approximable vectors in $\mathbb{R}^n$ for $n > 1$ is not absolutely winning.

In [Bengoechea et al. 2017] another strong variant of the winning property was given. We say that a set $E \subset \mathbb{R}^n$ is *isotropically winning* if for each $d \leq n$ and for each d-dimensional affine subspace $\mathcal{A} \subset \mathbb{R}^n$ the intersection $E \cap \mathcal{A}$ is $\frac{1}{2}$-winning for Schmidt's game considered as a game in $\mathcal{A}$. It is clear that an absolute winning set is isotropically winning for each $\alpha \leq \frac{1}{2}$.

3. Inhomogeneous approximations

The first important result on inhomogeneous approximations in the one-dimensional case is due to Khinchine [1926]. He proved that there exists an absolute constant γ such that for every $\theta \in \mathbb{R}$ there exists $\eta \in \mathbb{R}$ such that

$$\inf_{q \in \mathbb{Z}} q \|q\theta - \eta\| > \gamma.$$

Later (see [Khinchin 1937; 1948]) he proved that for given positive numbers $n, m \in \mathbb{Z}$ there exists a positive constant γ_{nm} such that for any $m \times n$ real matrix $\boldsymbol{\theta}$ there exists a vector $\boldsymbol{\eta} \in \mathbb{R}^n$ such that

$$\inf_{\boldsymbol{x} \in \mathbb{Z}^m \setminus \{0\}} (\|\boldsymbol{\theta}\boldsymbol{x} - \boldsymbol{\eta}\|_{\mathbb{Z}^n})^n \|\boldsymbol{x}\|^m > \gamma_{nm}$$

(here $\|\cdot\|_{\mathbb{Z}^n}$ stands for the distance to the nearest integral point in sup-norm). These results are presented in a wonderful book by Cassels [1957].

Jarník [1941], proved a generalization of this statement. Suppose $\psi(t)$ is a function decreasing to zero as $t \to +\infty$. Let $\rho(t)$ be the function inverse to the function $t \mapsto 1/\psi(t)$. Suppose that for all $t > 1$ one has $\psi_\theta(t) \leq \psi(t)$. Then there exists a vector $\boldsymbol{\eta} \in \mathbb{R}^n$ such that

$$\inf_{\boldsymbol{x} \in \mathbb{Z}^m \setminus \{0\}} (\|\boldsymbol{\theta}\boldsymbol{x} - \boldsymbol{\eta}\|_{\mathbb{Z}^n}) \cdot \rho(8m \cdot \|\boldsymbol{x}\|) > \gamma$$

with appropriate $\gamma = \gamma(n, m)$.

Denote by

$$\mathrm{Bad}_\theta = \big\{\alpha \in [0, 1) : \inf_{q\in\mathbb{N}} q \cdot \|q\theta - \alpha\| > 0\big\}.$$

It happens that the winning property of this inhomogeneous Diophantine set was considered quite recently. Tseng [2009] showed that Bad_θ is winning for all real numbers θ in classical Schmidt's sense. For the corresponding multidimensional sets

$$\mathrm{Bad}(n, m) = \big\{\boldsymbol{\theta} \in \mathrm{Mat}_{n\times m}(\mathbb{R}) : \inf_{q\in\mathbb{Z}^m_{\neq 0}} \max_{1\le i\le n} (|q|^{m/n}\|\boldsymbol{\theta}_i(q)\|) > 0\big\}.$$

the winning property is shown, for example, in [Einsiedler and Tseng 2011; Moshchevitin 2011]. In [Broderick et al. 2013] it was shown that the set $\mathrm{Bad}(n, m)$ is hyperplane absolutely winning. The methods used in [Broderick et al. 2013] come from [Broderick et al. 2011].

Further generalizations deal with the *twisted* sets

$$\mathrm{Bad}(i, j) = \big\{(\theta_1, \theta_2) \in \mathbb{R}^2 : \inf_{q\in\mathbb{N}} \max(q^i\|q\theta_1\|, q^j\|q\theta_2\|) > 0\big\},$$

where i, j are real positive numbers satisfying $i + j = 1$, introduced by Schmidt. In [An 2016] it was proved that $\mathrm{Bad}(i, j)$ is winning for the standard Schmidt game. In higher dimension, we fix an n-tuple $\boldsymbol{k} = (k_1, \dots, k_n)$ of real numbers satisfying

$$k_1, \dots, k_n > 0 \quad \text{and} \quad \sum_{i=1}^{n} k_i = 1, \tag{1}$$

and define

$$\mathrm{Bad}(\boldsymbol{k}, n, m) = \big\{\boldsymbol{\theta} \in \mathrm{Mat}_{n\times m}(\mathbb{R}) : \inf_{q\in\mathbb{Z}^m_{\neq 0}} \max_{1\le i\le n} (|q|^{mk_i}\|\boldsymbol{\theta}_i(q)\|) > 0\big\}.$$

Here, $|\cdot|$ denotes the supremum norm, $\boldsymbol{\theta} = (\boldsymbol{\theta}_{ij})$, and $\boldsymbol{\theta}_i(q)$ is the product of the i-th line of $\boldsymbol{\theta}$ with the vector q, i.e.,

$$\boldsymbol{\theta}_i(q) = \sum_{j=1}^{m} q_j\boldsymbol{\theta}_{ij}.$$

In the twisted setting, much less is known. In particular up to now the winning property of the set $\mathrm{Bad}(\boldsymbol{k}, n, m)$ in dimension greater that two is not proved.

Given $\boldsymbol{\theta} \in \mathrm{Mat}_{n\times m}(\mathbb{R})$, we define

$$\mathrm{Bad}_{\boldsymbol{\theta}}(\boldsymbol{k}, n, m) = \big\{x \in \mathbb{R}^n : \inf_{q\in\mathbb{Z}^m_{\neq 0}} \max_{1\le i\le n} (|q|^{mk_i}\|\boldsymbol{\theta}_i(q) - x_i\|) > 0\big\}.$$

Harrap and Moshchevitin [2017] showed that this set is winning provided that $\boldsymbol{\theta} \in \mathrm{Bad}(\boldsymbol{k}, n, m)$. In [Bengoechea et al. 2017] it was proved that if we suppose that $\boldsymbol{\theta} \in \mathrm{Bad}(\boldsymbol{k}, n, m)$, the set $\mathrm{Bad}_{\boldsymbol{\theta}}(\boldsymbol{k}, n, m)$ is isotropically winning.[1]

We should note that even in the case $n = 2, m = 1$ it is not known if the set $\mathrm{Bad}_{\boldsymbol{\theta}}(\boldsymbol{k}, 2, 1)$ is α-winning for some positive α without the condition $\boldsymbol{\theta} \in \mathrm{Bad}(\boldsymbol{k}, 2, 1)$.

[1] In fact, the approach from [Bengoechea et al. 2017] gives a little bit more. Instead of property that for any subspace $\mathcal{A}$ the intersection $E \cap \mathcal{A}$ is $\frac{1}{2}$-winning in $\mathcal{A}$, one can see that it is α-winning for all $\alpha \in \big(0, \frac{1}{2}\big]$. It is not completely clear for the author if these two properties are equivalent. (For a closely related problem, see [Dremov 2002].)

In this article we show that the condition $\boldsymbol{\theta}$ be from $\mathrm{Bad}(\boldsymbol{k}, n, m)$ is essential for the isotropically winning property, and prove the following theorem.

Theorem 3.1. *There exists a vector $\boldsymbol{\theta} = (\theta_1, \theta_2)$ such that*:

(1) $1, \theta_1, \theta_2$ *are linearly independent over* $\mathbb{Z}$.

(2) $\mathrm{Bad}_{\boldsymbol{\theta}} := \{(\eta_1, \eta_2) : \inf_{x\in\mathbb{N}} x^{1/2} \max_{i=1,2} \|x\theta_i - \eta_i\| > 0\}$ *is not isotropically winning.*

4. Some more remarks

In the sequel, $\boldsymbol{x} = (x_0, x_1, x_2)$ is a vector in $\mathbb{R}^3$, $|\cdot|$ stands for the Euclidean norm of the vector, and by $(\boldsymbol{w}, \boldsymbol{t})$ we denote the inner product of vectors $\boldsymbol{w}$ and $\boldsymbol{t}$.

The proof of Theorem 3.1 we will give in Section 6. There we will construct a special $\boldsymbol{\theta}$ and a one-dimensional affine subspace $\mathcal{P}$ such that $\boldsymbol{\theta} \in \mathcal{P}$ and for the segment $\mathcal{D} = \mathcal{P} \cap \{|\boldsymbol{z} - \boldsymbol{\theta}| \le 1\}$ one has $\mathcal{D} \cap \mathrm{Bad}_{\boldsymbol{\theta}} = \varnothing$. Moreover, given an arbitrary positive function $\omega(t)$ monotonically (slowly) increasing to infinity we can ensure that for all $\boldsymbol{\eta} = (\eta_1, \eta_2) \in \mathcal{D}$ there exist infinitely many $x \in \mathbb{Z}$ such that

$$\max_{i=1,2} \|x\theta_i - \eta_i\| < \frac{\omega(x)}{x}.$$

To explain the construction of the proof it is useful to consider the case when $\theta_1, \theta_2, 1$ are linearly dependent. This case we will discuss in Section 5.

Remark 4.1. From the result of the paper [Bengoechea et al. 2017] it follows that the vector $\boldsymbol{\theta}$ constructed in Theorem 3.1 does not belong to the set

$$\mathrm{Bad} = \big\{(\theta_1, \theta_2) \mid \inf_{x\in\mathbb{N}} x^{1/2} \max(\|\theta_1 x\|, \|\theta_2 x\|) > 0\big\}.$$

Remark 4.2. Let $\boldsymbol{\theta} = (a_1/q, a_2/q)$ be rational. Let $\boldsymbol{\eta} = (\eta_1, \eta_2) \notin \frac{1}{q} \cdot \mathbb{Z}^2$; then for any $x \in \mathbb{Z}$,

$$\max_{i=1,2} \left\| x\frac{a_i}{q} - \eta_i \right\| \ge \mathrm{dist}\big(\boldsymbol{\eta}, \tfrac{1}{q} \cdot \mathbb{Z}^2\big) > 0.$$

So the set

$$\mathcal{B} = \left\{ \boldsymbol{\eta} : \inf_{x\in\mathbb{Z}} \max_{i=1,2} \left\| x\frac{a_i}{q} - \eta_i \right\| > 0 \right\}$$

contains $\mathbb{R}^2 \setminus \frac{1}{q} \cdot \mathbb{Z}^2$ and is trivially winning. It is clear that for any one-dimensional affine subspace ℓ we have $\mathcal{B} \cap \ell \supset \big(\mathbb{R}^2 \setminus \frac{1}{q} \cdot \mathbb{Z}^2\big) \cap \ell$. So obviously $\mathcal{B} \cap \ell$ is also winning in ℓ.

5. Linearly dependent case

Let $1, \theta_1, \theta_2$ be linearly dependent and at least one of θ_j is irrational. This means that there exists $\boldsymbol{z} = (z_0, z_1, z_2) \in \mathbb{Z}^3$ such that $(\boldsymbol{z}, \boldsymbol{\theta}) = 0$. Let us consider the two-dimensional rational subspace

$$\pi = \{\boldsymbol{x} \in R^3 : (\boldsymbol{x}, \boldsymbol{z}) = 0\},$$

so $\boldsymbol{\theta} \in \pi$.

Let us define the one-dimensional subspace $\mathcal{P} = \{(x_1, x_2) : (1, x_1, x_2) \in \pi\} \subset \mathbb{R}^2$.

We will prove that there exists a constant γ such that for any $\eta = (\eta_1, \eta_2) \in \mathcal{P}$ the inequality

$$\max_{i=1,2} \|\theta_i x - \eta_i\| < \frac{\gamma}{x}$$

has infinitely many solutions in $x \in \mathbb{N}$. (This statement is similar to Chebyshev's theorem [Khinchin 1964, Theorem 24, Chapter 2].)

Denote by $\Lambda = \pi \cap \mathbb{Z}^3$ the integer lattice with the determinant $d := \det \Lambda = |z|$. Denote by $\{\boldsymbol{g}_\nu = (q_\nu, a_{1\nu}, a_{2\nu})\}_{\nu=1,2,3,\ldots} \subset \Lambda$ the sequence of the best approximations of $\boldsymbol{\theta}$ by the lattice Λ and the corresponding parallelograms

$$\Pi_\nu = \big\{\boldsymbol{x} = (x_0, x_1, x_2) \in \pi : 0 \le x_0 \le q_\nu,\ \operatorname{dist}(\boldsymbol{x}, l(\boldsymbol{\theta})) \le \operatorname{dist}(\boldsymbol{g}_{\nu-1}, l(\boldsymbol{\theta}))\big\},$$

which contains a fundamental domain of the two-dimensional Λ. Obviously, $\operatorname{vol} \Pi_\nu \le 4d$. So,

$$\operatorname{dist}(\boldsymbol{g}_{\nu-1}, l(\theta)) \ll \frac{d}{q_\nu}, \tag{2}$$

with an absolute constant in the sign $\ll$. It is clear that for any point $\boldsymbol{\eta} \in \pi$, the shift $\boldsymbol{\eta} + \Pi_\nu$ contains a point of Λ.

For any $\boldsymbol{\eta} = (\eta_1, \eta_2) \in \mathcal{P}$ and for any positive integer ν the planar domain $\bar{\boldsymbol{\eta}} + \Pi_\nu$, $\bar{\boldsymbol{\eta}} = (1, -\eta_1, -\eta_2)$ contains an integer point $\boldsymbol{y} = (x, y_1, y_2) \in \Lambda$.

It is clear that

$$1 \le x \le 1 + q_\nu \tag{3}$$

and

$$\max_{i=1,2} \|\theta_i x - \eta_i\| \ll \operatorname{dist}(\boldsymbol{y}, l(\boldsymbol{\theta}) + \bar{\boldsymbol{\eta}}) \ll \operatorname{dist}(l(\boldsymbol{\theta}), \boldsymbol{g}_{\nu-1}),$$

and by (2),

$$\max_{i=1,2} \|\theta_i x - \eta_i\| \ll \frac{d}{q_\nu}. \tag{4}$$

From (3), (4) it follows that the inequality

$$\max_{i=1,2} \|\theta_i x - \eta_i\| \ll \frac{d}{x}$$

has infinitely many solutions and everything is proved.

6. Inductive construction of integer points

Let $\omega(t)$ be arbitrary positive function monotonically (slowly) increasing to infinity. Here we describe the inductive construction of integer points $\boldsymbol{z}_\nu = (q_\nu, z_{1\nu}, z_{2\nu})$. The base of the induction process is trivial. One can take an arbitrary primitive pair of integer vectors that can be completed to a basis of $\mathbb{Z}^3$.

Suppose that we have two primitive integer vectors

$$\boldsymbol{z}_{\nu-1} = (q_{\nu-1}, z_{1\,\nu-1}, z_{2\,\nu-1}) \in \mathbb{Z}^3, \quad \boldsymbol{z}_\nu = (q_\nu, z_{1\,\nu}, z_{2\,\nu}) \in \mathbb{Z}^3.$$

Now we explain how to construct the next integer vector $\boldsymbol{z}_{\nu+1}$.

We consider the two-dimensional subspace

$$\pi_\nu = \langle \boldsymbol{z}_{\nu-1}, \boldsymbol{z}_\nu \rangle_{\mathbb{R}}.$$

The pair of vectors $z_{\nu-1}$ and z_ν is primitive, so the lattice spanned by them is

$$\Lambda_\nu := \langle z_{\nu-1}, z_\nu\rangle_{\mathbb{Z}} = \pi_\nu \cap \mathbb{Z}^3.$$

By $d_\nu = \det \Lambda_\nu$ we denote the two-dimensional fundamental volume of the lattice Λ_ν. Now we define the vector $\boldsymbol{n}_\nu = (n_{0\nu}, n_{1\nu}, n_{2\nu}) \in \mathbb{R}^3$ from the conditions

$$\pi_\nu = \{\boldsymbol{x} \in \mathbb{R}^3 : (\boldsymbol{x}, \boldsymbol{n}_\nu) = 0\}, \quad |\boldsymbol{n}_\nu| = 1.$$

Put

$$\sigma_\nu = \operatorname{dist}(z_{\nu-1}, l(z_\nu)). \tag{5}$$

Obviously, $|z_\nu| \asymp q_\nu$ and

$$\sigma_\nu \asymp \frac{d_\nu}{q_\nu}. \tag{6}$$

We define a vector $\boldsymbol{e}_\nu$ from the conditions

$$\boldsymbol{e}_\nu \in \pi_\nu, \quad |\boldsymbol{e}_\nu| = 1, \quad (\boldsymbol{e}_\nu, z_\nu) = 0, \tag{7}$$

so $\boldsymbol{e}_\nu$ is parallel to π_ν and orthogonal to z_ν.

Define the rectangle

$$\Pi_\nu = \big\{\boldsymbol{x} = (x_0, x_1, x_2) : \boldsymbol{x} = t z_\nu + r\boldsymbol{e}_\nu,\ 0 \le t \le |z_\nu|,\ |r| \le \sigma_\nu\big\}.$$

It is clear that rectangle $\Pi_\nu \subset \pi_\nu$ contains a fundamental domain of the lattice Λ_ν. We need two axillary vectors z_ν^a and z_ν^b defined as

$$z_\nu^a = z_\nu + a_\nu \boldsymbol{e}_\nu, \quad z_\nu^b = z_\nu^a + b_\nu \boldsymbol{n}_\nu,$$

where positive a_ν is chosen in such a way that

$$a_\nu d_\nu^2 \le \nu^{-1}\omega\left(\frac{q_\nu^2}{d_\nu^2}\cdot\frac{1}{a_\nu}\right) \tag{8}$$

and

$$b_\nu = a_\nu \min\left(1, \frac{d_\nu}{q_\nu}\right). \tag{9}$$

From the construction, it follows that

$$|z_\nu^a| \asymp |z_\nu^b| \asymp |z_\nu| \asymp q_\nu. \tag{10}$$

The integer lattice $\mathbb{Z}^3$ splits into levels with respect to the two-dimensional sublattice Λ_ν in such a way that

$$\mathbb{Z}^3 = \bigsqcup_{i\in\mathbb{Z}} \Lambda_{\nu,i},$$

where $\Lambda_{\nu,j} = \Lambda_\nu + jz'$, $j \in \mathbb{Z}$ and integer vector z' completes the couple $z_{\nu-1}$, z_ν to the basis in $\mathbb{Z}^3$. We consider the affine subspace $\pi_\nu^1 = \pi_\nu + z' \supset \Lambda_{\nu,1}$, which is parallel to π_ν. It is clear that $\operatorname{dist}(\pi_\nu, \pi_\nu^1) = 1/d_\nu$.

We need to determine the next integer point $z_{\nu+1}$. Denote by $\mathfrak{P}$ the central projection with center 0 onto the affine subspace π_ν^1. We consider the triangle Δ with vertices z_ν, z_ν^a, z_ν^b and its image $\mathfrak{P}\Delta$ under

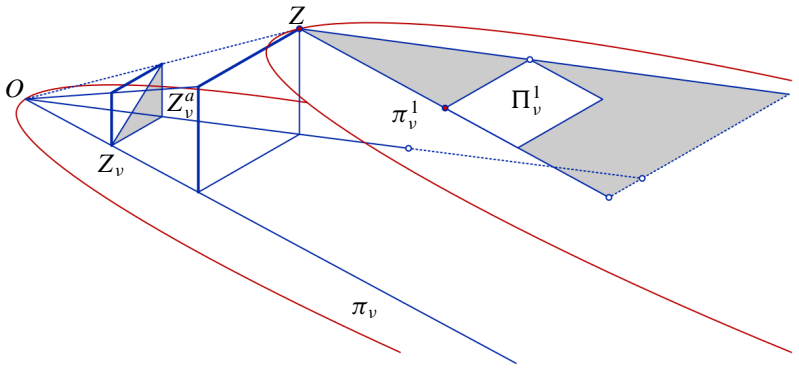


Figure 1. The central projection $\mathfrak{P}$.

the projection $\mathfrak{P}$ (Figure 1).

Define

$$\boldsymbol{Z} = \mathfrak{P} z_\nu^b. \tag{11}$$

One can see that

$$|\boldsymbol{Z}| \asymp \frac{q_\nu}{d_\nu b_\nu}. \tag{12}$$

Define rays

$$\mathcal{R}_1 = \{z = \boldsymbol{Z} + t z_\nu : t \geq 0\} \quad \text{and} \quad \mathcal{R}_2 = \{z = \boldsymbol{Z} + t z_\nu^a : t \geq 0\}.$$

It is clear that $\mathcal{R}_1 \cap \mathcal{R}_2 = \{\boldsymbol{Z}\}$ and $\mathcal{R}_1, \mathcal{R}_2 \subset \pi_\nu^1$. Moreover, the whole convex angle bounded by rays $\mathcal{R}_1, \mathcal{R}_2$ form the image of the triangle Δ under the projection $\mathfrak{P}$:

$$\mathfrak{P}\Delta = \operatorname{conv}(\mathcal{R}_1 \cup \mathcal{R}_2).$$

The affine subspace π_ν^1 contains the affine lattice $\Lambda_\nu^1 = \Lambda_\nu + z'$ which is congruent to the lattice Λ_ν. Thus, for any $\boldsymbol{\zeta} \in \pi_\nu^1$, the shift $\Pi_\nu + \boldsymbol{\zeta}$ contains an integer point from Λ_ν^1.

Put

$$\tau_\nu = \frac{2\sigma_\nu |z_\nu|}{a_\nu}. \tag{13}$$

Consider the point

$$\boldsymbol{\zeta}_\nu = \boldsymbol{Z} + \tau_\nu z_\nu + \sigma_\nu \boldsymbol{e}_\nu \in \pi_\nu^1,$$

and the rectangle

$$\Pi_\nu^1 = \Pi_\nu + \boldsymbol{\zeta}_\nu \subset \pi_\nu^1.$$

It is clear that

$$\Pi_\nu^1 \subset \mathfrak{P}\Delta$$

(here $\boldsymbol{Z}$ was defined in (11), $\boldsymbol{e}_\nu$ was defined in (7), and the parameters σ_ν, τ_ν come from (5) and (13)).

Now we take the integer point

$$z_{\nu+1} = (q_{\nu+1}, z_{1\,\nu+1}, z_{2\,\nu+1}) \in \Lambda_\nu^1 \cap \Pi_\nu^1.$$

From the construction it follows that

$$q_{\nu+1} \asymp |z_{\nu+1}| \asymp |z| + \tau_\nu |z_\nu| + |z_\nu| \asymp q_\nu \left(1 + \frac{1}{d_\nu b_\nu} + \frac{\sigma_\nu}{a_\nu}\right) \asymp q_\nu \left(1 + \frac{1}{d_\nu b_\nu}\right) + \frac{d_\nu}{a_\nu} \asymp \frac{q_\nu}{d_\nu b_\nu}.$$

(Here we use (6), (9), (10), (12), and (13).) From (9) we see that

$$q_{\nu+1} \gg \left(\frac{q_\nu}{d_\nu}\right)^2 \frac{1}{a_\nu}. \tag{14}$$

Now we are able to define the next two-dimensional lattice

$$\Lambda_{\nu+1} = \langle z_\nu, z_{\nu+1} \rangle_{\mathbb{Z}}.$$

Let $d_{\nu+1}$ be its fundamental volume. We will estimate the value of $d_{\nu+1}$ taking into account (9) as

$$d_{\nu+1} \ll q_\nu \cdot \operatorname{dist}(z_{\nu+1}, l(z_\nu)) \ll \frac{q_\nu}{d_\nu} \cdot \frac{a_\nu}{b_\nu} \ll \left(\frac{q_\nu}{d_\nu}\right)^2 \ll q_\nu^2. \tag{15}$$

From (14) and (15), we deduce that

$$d_{\nu+1} \ll a_\nu d_\nu^2 q_{\nu+1}.$$

By the choice of a_ν (by formula (8)) we have

$$d_{\nu+1} \le \frac{\omega(q_{\nu+1})}{\nu}. \tag{16}$$

7. The vector $\boldsymbol{\theta}$

Now we define

$$\boldsymbol{\theta}_\nu = (\theta_{1\nu}, \theta_{2\nu}), \quad \theta_{j\nu} = \frac{q_{j\nu}}{q_\nu}.$$

We consider the angles between the successive vectors $\boldsymbol{n}_\nu$ and $\boldsymbol{n}_{\nu+1}$:

$$\alpha_\nu = \operatorname{angle}(\boldsymbol{n}_\nu, \boldsymbol{n}_{\nu+1}) \asymp \tan \operatorname{angle}(\boldsymbol{n}_\nu, \boldsymbol{n}_{\nu+1}).$$

Since $z_{\nu+1} \in \mathfrak{P}\Delta$ (see Figure 2), we have

$$\tan \operatorname{angle}(\boldsymbol{n}_\nu, \boldsymbol{n}_{\nu+1}) \le \frac{b_\nu}{a_\nu},$$

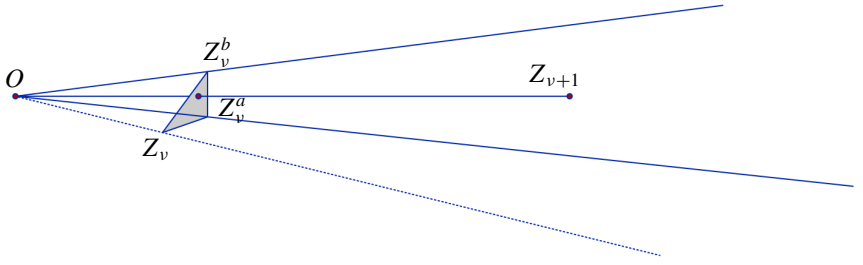


Figure 2. The vector $z_{\nu+1}$ intersects the interior of the triangle $\Delta = z_\nu z_\nu^a z_\nu^b$.

and so

$$\alpha_\nu \ll \frac{b_\nu}{a_\nu}. \tag{17}$$

As $z_{\nu+1} \in \mathfrak{P}\Delta$, we have

$$|\boldsymbol{\theta}_\nu - \boldsymbol{\theta}_{\nu+1}| \ll \frac{\sqrt{a_\nu^2 + b_\nu^2}}{q_\nu} \ll \frac{a_\nu}{q_\nu} \tag{18}$$

by the same argument. There exist limits

$$\lim_{\nu\to\infty} \boldsymbol{\theta}_\nu = \boldsymbol{\theta} = (\theta_1, \theta_2) \quad \text{and} \quad \lim_{\nu\to\infty} \boldsymbol{n}_\nu = \boldsymbol{n},$$

and from (17) and (18) we deduce that

$$0 < |\boldsymbol{\theta} - \boldsymbol{\theta}_\nu| \ll \frac{a_\nu}{q_\nu} \tag{19}$$

and

$$\operatorname{angle}(\boldsymbol{n}, \boldsymbol{n}_\nu) \ll \frac{b_\nu}{a_\nu}. \tag{20}$$

It is clear that $\boldsymbol{\theta} \notin \mathbb{Q}^2$. A slight modification[2] of the procedure of choosing vectors z_ν ensures the condition that $1, \theta_1, \theta_2$ are linearly independent over $\mathbb{Z}$. Define $\pi = \{\boldsymbol{x} \in \mathbb{R}^3 : (\boldsymbol{x}, \boldsymbol{n}) = 0\}$. Then $\boldsymbol{\theta} \in \pi$ by continuity and we can assume that $\boldsymbol{n} \notin \mathbb{Q}^3$.

8. Winning property

Consider the one-dimensional affine subspaces

$$\mathcal{P}_\nu = \{(x_1, x_2) \in \mathbb{R}^2 : (1, x_1, x_2) \in \pi_\nu\} \subset \mathbb{R}^2$$

and

$$\mathcal{P} = \{(x_1, x_2) \in \mathbb{R}^2 : (1, x_1, x_2) \in \pi\} \subset \mathbb{R}^2,$$

where π was defined at the end of the previous section. Let

$$B_1(\boldsymbol{\theta}) = \{\boldsymbol{\xi} \in \mathbb{R}^2 : \operatorname{dist}(\boldsymbol{\xi}, \boldsymbol{\theta}) < 1\}.$$

We will show that for any $\boldsymbol{\eta} = (\eta_1, \eta_2) \in \mathcal{P} \cap B_1(\boldsymbol{\theta})$ there exists infinitely many solutions of the inequality

$$\max_{i=1,2} \|\theta_i x - \eta_i\| < \frac{\omega(x)}{x}$$

in integers x. Denote by $\boldsymbol{\eta}_\nu = (\eta_{1\nu}, \eta_{2\nu})$ the orthogonal projection of $\boldsymbol{\eta}$ onto $\mathcal{P}_\nu$. From (20) we see that

$$|\boldsymbol{\eta} - \boldsymbol{\eta}_\nu| \ll \frac{b_\nu}{a_\nu}. \tag{21}$$

[2] A similar procedure was explained in [Moshchevitin 2012]. There, the author provides the linear independence of coordinates of the limit vector by "going away from all rational subspaces" (the beginning of the proof of Theorem 1 in the case $k = 1$, p. 132 and the beginning of §5, p. 146).

For any $\boldsymbol{\eta}_\nu = (\eta_{1\nu}, \eta_{2\nu}) \in \mathcal{P}_\nu$ the planar domain $\bar{\boldsymbol{\eta}}_\nu + \Pi_\nu$, $\bar{\boldsymbol{\eta}}_\nu = (1, -\eta_{1\nu}, -\eta_{2\nu})$ contains an integer point $\boldsymbol{y}_\nu = (x_\nu, y_{1\nu}, y_{2\nu}) \in \Lambda_\nu$. It is clear that

$$|x_\nu| \ll q_\nu \tag{22}$$

and

$$\max_{i=1,2} |\theta_{i\nu} x_\nu - \eta_{i\nu} - y_{i\nu}| \ll \frac{d_\nu}{q_\nu}. \tag{23}$$

By (19), (21), (22), and (23) we have

$$\max_{i=1,2} \|\theta_i x_\nu - \eta_i\| \le |x_\nu| \max_{i=1,2} |\theta_i - \theta_{i\nu}| + \max_{i=1,2} \|\theta_{i\nu} x_\nu - \eta_{i\nu}\| + \max_{i=1,2} |\eta_i - \eta_{i\nu}| \ll a_\nu + \frac{d_\nu}{q_\nu} + \frac{b_\nu}{a_\nu} \ll \frac{d_\nu}{q_\nu}.$$

In the last inequality we use (9). By (16) we have

$$\max_{i=1,2} \|\theta_i x_\nu - \eta_i\| \le \frac{\omega(q_\nu)}{q_\nu}$$

for large ν. As $\bar{\boldsymbol{\eta}} \in \pi$ and $\boldsymbol{y}_\nu \in \pi_\nu$, $\max_{i=1,2} \|\theta_i x_\nu - \eta_i\| \ne 0$ infinitely often (in fact for all large ν).

References

[An 2016] J. An, "2-dimensional badly approximable vectors and Schmidt's game", *Duke Math. J.* **165**:2 (2016), 267–284. MR Zbl

[Bengoechea et al. 2017] P. Bengoechea, N. Moshchevitin, and N. Stepanova, "A note on badly approximable linear forms on manifolds", *Mathematika* **63**:2 (2017), 587–601. MR Zbl

[Broderick et al. 2011] R. Broderick, L. Fishman, and D. Kleinbock, "Schmidt's game, fractals, and orbits of toral endomorphisms", *Ergodic Theory Dynam. Systems* **31**:4 (2011), 1095–1107. MR Zbl

[Broderick et al. 2013] R. Broderick, L. Fishman, and D. Simmons, "Badly approximable systems of affine forms and incompressibility on fractals", *J. Number Theory* **133**:7 (2013), 2186–2205. MR Zbl

[Cassels 1957] J. W. S. Cassels, *An introduction to Diophantine approximation*, Cambridge Tracts in Mathematics and Mathematical Physics **45**, Cambridge University Press, 1957. MR Zbl

[Dremov 2002] V. A. Dremov, "On domains of (α, β)-winnability", *Dokl. Akad. Nauk* **384**:3 (2002), 304–307. In Russian; translated in *Dokl. Math.* **65**:3 (2002), 365–368. MR Zbl

[Einsiedler and Tseng 2011] M. Einsiedler and J. Tseng, "Badly approximable systems of affine forms, fractals, and Schmidt games", *J. Reine Angew. Math.* **660** (2011), 83–97. MR Zbl

[Fishman et al. 2013] L. Fishman, D. S. Simmons, and M. Urbański, "Diophantine approximation and the geometry of limit sets in Gromov hyperbolic metric spaces", 2013. arXiv

[Harrap and Moshchevitin 2017] S. Harrap and N. Moshchevitin, "A note on weighted badly approximable linear forms", *Glasg. Math. J.* **59**:2 (2017), 349–357. MR Zbl

[Jarník 1941] V. Jarník, "O lineárních nehomogenních diofantických aproximacích", *Rozpr. II. Třídy České Akad.* **51**:29 (1941), 1–21. Translated in French as "Sur les approximations diophantiques linéaires non homogènes", *Acad. Tchèque Sci. Bull. Int. Cl. Sci. Math. Nat.* **47** (1946), 145-160. MR Zbl

[Khinchin 1926] A. Khintchine, "Über eine Klasse linearer diophantischer Approximationen", *Rend. Circ. Mat. Palermo* **50** (1926), 170–195. Zbl

[Khinchin 1937] A. Khintchine, "Über die angenäherte Auflösung linearer Gleichungen in ganzen Zahlen", *Acta Arith.* **2** (1937), 161–172. Zbl

[Khinchin 1948] A. Y. Khinchin, "Regular systems of linear equations and a general problem of Chebyshev", *Izvestiya Akad. Nauk SSSR. Ser. Mat.* **12** (1948), 249–258. In Russian. MR

[Khinchin 1964] A. Y. Khinchin, *Continued fractions*, The University of Chicago Press, 1964. MR Zbl

[Kleinbock and Weiss 2010] D. Kleinbock and B. Weiss, "Modified Schmidt games and Diophantine approximation with weights", *Adv. Math.* **223**:4 (2010), 1276–1298. MR Zbl

[McMullen 2010] C. T. McMullen, "Winning sets, quasiconformal maps and Diophantine approximation", *Geom. Funct. Anal.* **20**:3 (2010), 726–740. MR Zbl

[Moshchevitin 2011] N. G. Moshchevitin, "A note on badly approximable affine forms and winning sets", *Mosc. Math. J.* **11**:1 (2011), 129–137. MR Zbl

[Moshchevitin 2012] N. G. Moshchevitin, "Proof of W. M. Schmidt's conjecture concerning successive minima of a lattice", *J. Lond. Math. Soc.* (2) **86**:1 (2012), 129–151. MR Zbl

[Schmidt 1966] W. M. Schmidt, "On badly approximable numbers and certain games", *Trans. Amer. Math. Soc.* **123** (1966), 178–199. MR Zbl

[Schmidt 1969] W. M. Schmidt, "Badly approximable systems of linear forms", *J. Number Theory* **1** (1969), 139–154. MR Zbl

[Schmidt 1980] W. M. Schmidt, *Diophantine approximation*, Lecture Notes in Mathematics **785**, Springer, 1980. MR Zbl

[Tseng 2009] J. Tseng, "Badly approximable affine forms and Schmidt games", *J. Number Theory* **129**:12 (2009), 3020–3025. MR Zbl

Received 1 Dec 2017.

NATALIA DYAKOVA:

natalia.stepanova.msu@gmail.com
Department of Mathematics and Mechanics, Moscow State University, Moscow, Russia

msp

Moscow Journal of Combinatorics and Number Theory
Vol. 8, No. 1, 2019
dx.doi.org/10.2140/moscow.2019.8.15

Some remarks on the asymmetric sum-product phenomenon

Ilya D. Shkredov

Using some new observations connected to higher energies, we obtain quantitative lower bounds on $\max\{|AB|, |A+C|\}$ and $\max\{|AB|, |(A+\alpha)C|\}$, where $\alpha \neq 0$, in the regime when the sizes of the finite subsets A, B, C of a field differ significantly.

1. Introduction

Let p be a prime number and $A, B \subset \mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ be finite sets. Define the *sum set*, the *difference set*, the *product set*, and the *quotient set* of A and B as

$$A+B := \{a+b : a \in A, b \in B\}, \quad A-B := \{a-b : a \in A, b \in B\},$$
$$AB := \{ab : a \in A, b \in B\}, \qquad A/B := \{a/b : a \in A, b \in B, b \neq 0\}.$$

One of the central problems in arithmetic combinatorics [Tao and Vu 2006] is the *sum-product problem*, which asks for estimates of the form

$$\max\{|A+A|, |AA|\} \geq |A|^{1+c} \tag{1}$$

for some positive c. This question was originally posed by Erdős and Szemerédi [1983] for finite sets of integers; they conjectured that (1) holds for all $c < 1$. The sum-product problem has since been studied over a variety of fields and rings; see, e.g., [Bourgain 2003; 2005b; 2007, Bush and Croot 2014; Bourgain et al. 2004; Erdős and Szemerédi 1983; Tao and Vu 2006]. We focus on the case of $\mathbb{F}_p$ (and sometimes consider $\mathbb{R}$), where the first estimate of the form (1) was proved by Bourgain, Katz, and Tao [Bourgain et al. 2004]. At the moment the best results in this direction are contained in [Roche-Newton et al. 2016; Konyagin and Shkredov 2016].

In this article we study an asymmetric variant of the sum-product question, in the spirit of the fundamental paper [Bourgain 2005c]: namely, sum-product theorems in $\mathbb{F}_p$ for sets of distinct sizes. We recall two results from that paper:

Theorem 1. *Given $0 < \varepsilon < \frac{1}{10}$, there is $\delta > 0$ such that the following holds. Let $A \subset \mathbb{F}_p$ be such that $p^{\varepsilon} < |A| < p^{1-\varepsilon}$. Then either*

$$|AB| > p^{\delta}|A| \quad \textit{for all } B \subset \mathbb{F}_p \textit{ with } |B| > p^{\varepsilon}$$

or

$$|A+C| > p^{\delta}|A| \quad \textit{for all } C \subset \mathbb{F}_p \textit{ with } |C| > p^{\varepsilon}.$$

MSC2010: 11B30, 11P70.
Keywords: sum-product, expanders, exponential sums.

Theorem 2. *Given $0 < \varepsilon < \frac{1}{10}$, there is $\delta > 0$ such that the following holds. Let $A \subset \mathbb{F}_p$ be such that $p^{\varepsilon} < |A| < p^{1-\varepsilon}$. Then for any $x \neq 0$ either*

$$|AB| > p^{\delta}|A| \quad \text{for all } B \subset \mathbb{F}_p \text{ with } |B| > p^{\varepsilon}$$

or

$$|(A+x)C| > p^{\delta}|A| \quad \text{for all } C \subset \mathbb{F}_p \text{ with } |C| > p^{\varepsilon}.$$

Theorems 1 and 2 were derived in [Bourgain 2005c] from the following result from [Bourgain 2005a]. Given a set $A \subseteq \mathbb{F}_p$ denote by $\mathsf{T}_k^+(A) := |\{(a_1,\dots,a_k,a_1',\dots,a_k') \in A^{2k} : a_1+\dots+a_k = a_1'+\dots+a_k'\}|$. We write $\mathsf{E}^+(A)$ for $\mathsf{T}_2^+(A)$.

Theorem 3. *For a positive integer Q, there are a positive integer k and a real $\tau > 0$ such that if $H \subseteq \mathbb{F}_p^*$ and $|HH| < |H|^{1+\tau}$, then*

$$\mathsf{T}_k^+(H) < |H|^{2k}(p^{-1+1/Q} + c_Q|H|^{-Q}),$$

where $c_Q > 0$ depends on Q only.

The aim of this paper is to obtain explicit bounds in the theorems above. Our arguments are different and more elementary than those of [Bourgain 2005c; Bourgain et al. 2006; Garaev 2010]. In the proof we almost do not use the Fourier approach and that is why we do not need lower bounds for sizes of A, B, C in terms of the characteristic p, but, of course, these sets must be comparable somehow. Another difference between this article and [Bourgain 2005c] is that our arguments work in $\mathbb{R}$ as well.

We now formulate our variants of Theorems 1 and 2 (see also Corollary 33). One can show that Theorem 4 implies Theorems 1 and 2 if $|A| < p^{1/2-\varepsilon}$; see Remark 36.

Theorem 4. *Let $A, B, C \subseteq \mathbb{F}_p$ be arbitrary sets, and $k \geq 1$ be such that $|A||B|^{1+\frac{k+1}{2(k+4)}2^{-k}} \leq p$ and*

$$|B|^{\frac{k}{8}+\frac{1}{2(k+4)}} \geq |A| \cdot C_*^{(k+4)/4} \log^k(|A||B|), \tag{2}$$

where $C_* > 0$ is an absolute constant. Then

$$\max\{|AB|, |A+C|\} \geq 2^{-3}|A| \cdot \min\{|C|, |B|^{\frac{1}{2(k+4)}2^{-k}}\}, \tag{3}$$

and for any $\alpha \neq 0$

$$\max\{|AB|, |(A+\alpha)C|\} \geq 2^{-3}|A| \cdot \min\{|C|, |B|^{\frac{1}{2(k+4)}2^{-k}}\}. \tag{4}$$

Actually, we prove that the lower bounds for $|A+C|$, $|(A+\alpha)C|$ in (3), (4) could be replaced by similar upper bounds for the energies $\mathsf{E}^+(A,C)$, $\mathsf{E}^{\times}(A+\alpha,C)$; see the second part of Corollary 33. We call Theorem 4 an asymmetric sum-product result because A can be much larger than B and C (say, $|A| > (|B||C|)^{100}$) in contrast with the usual quadratic restrictions which follow from the classical Szemerédi–Trotter theorem; see [Szemerédi and Trotter 1983; Tao and Vu 2006] for the real setting and see [Bourgain et al. 2004; Garaev 2010; Rudnev 2017b] for prime fields. On the other hand, the roles of B, C are not symmetric as well. The thing is that the method of the proof intensively uses the fact that if $|AB|$ is small comparable to $|A|$, then, roughly speaking, for any integer k, the size of $(kA)B$ is small comparable to kA, roughly speaking (rigorous formulation can be found in Section 5). Of course this observation is not true in any sense if we replace $\times$ to $+$ and vice versa.

Also, we obtain a "quantitative" version of Theorem 3.

Theorem 5. *Let $A, B \subseteq \mathbb{F}_p$ be sets, $M \geq 1$ be a real number and $|AB| \leq M|A|$. For any $k \geq 2$ such that $2^{16k} M^{2^{k+1}} C_*^2 \log^8 |A| \leq |B|$, one has*

$$\mathsf{T}^+_{2^k}(A) \leq$$
$$2^{4k+6} C_* \log^4 |A| \cdot \frac{M^{2^k} |A|^{2^{k+1}}}{p} + 16^{k^2} M^{2^{k+1}} C_*^{k-1} \log^{4(k-1)} |A| \cdot |A|^{2^{k+1}-4} |B|^{-\frac{k-1}{2}} \mathsf{E}^+(A). \quad (5)$$

Here, $C_* > 0$ is an absolute constant.

As a by-product, we obtain the best constants in the problem of estimating the exponential sums over multiplicative subgroups [Bourgain 2005a; Garaev 2010] (see Corollary 16 below) and relatively good bounds in the question of basis properties of multiplicative subgroups [Glibichuk and Konyagin 2007]. Also, we find a wide series of "superquadratic expanders in $\mathbb{R}$" [Balog et al. 2017] with four variables; see Corollary 35.

In contrast to [Bourgain 2005c], we prove Theorem 4 and Theorem 5 independently. We realize that Theorem 4 is equivalent to estimating energies of another sort, namely,

$$\mathsf{E}^+_k(A) := \big|\{(a_1, \dots, a_k, a'_1, \dots, a'_k) \in A^{2k} : a_1 - a'_1 = \dots = a_k - a'_k\}\big|$$

(see the definitions in Section 2). Thus, a new feature of this paper is an upper bound for $\mathsf{E}^+_k(A)$ for sets A with $|AB| \ll |A|$ for some large B; see Theorem 27 below. Such an upper bound can be of independent interest. Let us formulate our result about $\mathsf{E}^+_k(A)$.

Theorem 6. *Let $A, B \subseteq \mathbb{F}_p$ be two sets, $k \geq 0$ be an integer, and put $M := |AB^{k+1}|/|A|$. Then for any $k \geq 0$ such that*

$$|B|^{k/8+1/2} \geq |A| \cdot M^{2^k+1} 2^{3k+1} C_*^{(k+4)/4} \log^k |AB^k|,$$

where $C_* > 0$ is an absolute constant, we have

$$\mathsf{E}^+_{2^{k+1}}(A) \leq 2|AB^k|^{2^{k+1}}. \quad (6)$$

Our approach develops the ideas from [Bourgain 2005c; Shkredov 2014] (see especially Section 4 there) and uses several sum-product observations of course. We avoid repeating Bourgain's combinatorial arguments (although we use a similar inductive proof strategy) but the method relies on recent geometrical sum-product bounds from [Rudnev 2017b] and further papers such as [Yazici et al. 2017; Murphy et al. 2017; Roche-Newton et al. 2016; Shkredov 2017]. In some sense we introduce a new approach of estimating moments $\mathrm{M}_k(f)$ (e.g., $\mathsf{T}^+_k(H)$ in Theorem 3 or $\mathsf{E}^+_k(A)$ in Theorem 6) of some specific functions f: instead of calculating $\mathrm{M}_k(f)$ in terms of suitable norms of f, we compare $\mathrm{M}_k(f)$ and $\mathrm{M}_{k/2}(f)$. If $\mathrm{M}_k(f)$ is much less than $\mathrm{M}_{k/2}(f)$, then we use induction, and if not, then thanks some special nature of the function f, we derive from this fact that the additive energy E^+ of a level set of f is huge and it gives a contradiction. Clearly, this process can be applied at most $O(\log k)$ number of times and that is why we usually have logarithmic savings (compare the index in $\mathsf{T}^+_{2^k}(A)$ and the gain $|B|^{-(k-1)/2}$ in estimate (5), say).

The paper is organized as follows. Section 2 contains all required definitions. In Section 3 we give a list of the results, which will be further used in the text. In Section 4, we consider a particular case

of multiplicative subgroups Γ and obtain an upper estimate for $\mathsf{T}_k^+(\Gamma)$. This technique is developed in Section 5 although we avoid using the Fourier approach as was done in [Bourgain 2005c] and in the previous Section 4. Section 5 contains all main Theorems 4–6.

2. Notation

In this paper p is an odd prime number, $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$, and $\mathbb{F}_p^* = \mathbb{F}_p \setminus \{0\}$. We denote the Fourier transform of a function $f : \mathbb{F}_p \to \mathbb{C}$ by $\hat{f}$,

$$\hat{f}(\xi) = \sum_{x\in\mathbb{F}_p} f(x)e(-\xi\cdot x), \tag{7}$$

where $e(x) = e^{2\pi i x/p}$. We rely on the following basic identities. The first one is called the Plancherel formula and its particular case $f = g$ is called the Parseval identity:

$$\sum_{x\in\mathbb{F}_p} f(x)\overline{g(x)} = \frac{1}{p}\sum_{\xi\in\mathbb{F}_p} \hat{f}(\xi)\overline{\hat{g}(\xi)}. \tag{8}$$

A particular case of (8) is

$$\sum_{y\in\mathbb{F}_p}\left|\sum_{x\in\mathbb{F}_p} f(x)g(y-x)\right|^2 = \frac{1}{p}\sum_{\xi\in\mathbb{F}_p} |\hat{f}(\xi)|^2|\hat{g}(\xi)|^2, \tag{9}$$

and the formula

$$f(x) = \frac{1}{p}\sum_{\xi\in\mathbb{F}_p} \hat{f}(\xi)e(\xi\cdot x) \tag{10}$$

is called the inversion formula. Further let $f, g : \mathbb{F}_p \to \mathbb{C}$ be two functions. Put

$$(f*g)(x) := \sum_{y\in\mathbb{F}_p} f(y)g(x-y) \quad \text{and} \quad (f\circ g)(x) := \sum_{y\in\mathbb{F}_p} \overline{f(y)}g(y+x). \tag{11}$$

Then

$$\widehat{f*g} = \hat{f}\hat{g} \quad \text{and} \quad \widehat{f\circ g} = \bar{\hat{f}}\hat{g}. \tag{12}$$

Put $\mathsf{E}^+(A, B)$ for the *common additive energy* of two sets $A, B \subseteq \mathbb{F}_p$ (see, e.g., [Tao and Vu 2006]); that is,

$$\mathsf{E}^+(A, B) = \big|\{(a_1, a_2, b_1, b_2) \in A\times A\times B\times B : a_1 + b_1 = a_2 + b_2\}\big|.$$

If $A = B$ we simply write $\mathsf{E}^+(A)$ instead of $\mathsf{E}^+(A, A)$ and $\mathsf{E}^+(A)$ is called the *additive energy* in this case. Clearly,

$$\mathsf{E}^+(A, B) = \sum_x (A*B)(x)^2 = \sum_x (A\circ B)(x)^2 = \sum_x (A\circ A)(x)(B\circ B)(x)$$

and by (9),

$$\mathsf{E}(A, B) = \frac{1}{p}\sum_{\xi} |\hat{A}(\xi)|^2|\hat{B}(\xi)|^2. \tag{13}$$

Also, notice that

$$\mathsf{E}^+(A, B) \le \min\{|A|^2|B|, |B|^2|A|, |A|^{3/2}|B|^{3/2}\}. \tag{14}$$

Sometimes we write $\mathsf{E}^+(f_1, f_2, f_3, f_4)$ for the additive energy of four real functions, namely,

$$\mathsf{E}^+(f_1, f_2, f_3, f_4) = \sum_{x,y,z} f_1(x) f_2(y) f_3(x+z) f_4(y+z).$$

It can be shown using the Hölder inequality (see, e.g., [Tao and Vu 2006]) that

$$\mathsf{E}^+(f_1, f_2, f_3, f_4) \le (\mathsf{E}^+(f_1)\mathsf{E}^+(f_1)\mathsf{E}^+(f_1)\mathsf{E}^+(f_1))^{1/4}. \tag{15}$$

In the same way define the *common multiplicative energy* of two sets $A, B \subseteq \mathbb{F}_p$:

$$\mathsf{E}^\times(A, B) = \big|\{(a_1, a_2, b_1, b_2) \in A \times A \times B \times B : a_1 b_1 = a_2 b_2\}\big|.$$

Certainly, the multiplicative energy $\mathsf{E}^\times(A, B)$ can be expressed in terms of multiplicative convolutions similar to (11).

Sometimes we use representation function notations like $r_{AB}(x)$ or $r_{A+B}(x)$, which counts the number of ways $x \in \mathbb{F}_p$ can be expressed as a product ab or a sum $a + b$ with $a \in A$, $b \in B$, respectively. For example, $|A| = r_{A-A}(0)$ and $\mathsf{E}^+(A) = r_{A+A-A-A}(0) = \sum_x r^2_{A+A}(x) = \sum_x r^2_{A-A}(x)$. In this paper, we use the same letter to denote a set $A \subseteq \mathbb{F}_p$ and its characteristic function $A : \mathbb{F}_p \to \{0, 1\}$. Thus, $r_{A+B}(x) = (A * B)(x)$, say.

Now consider two families of higher energies. Firstly, let

$$\mathsf{T}^+_k(A) := \big|\{(a_1, \dots, a_k, a'_1, \dots, a'_k) \in A^{2k} : a_1 + \dots + a_k = a'_1 + \dots + a'_k\}\big| = \frac{1}{p} \sum_\xi |\hat{A}(\xi)|^{2k}. \tag{16}$$

It is useful to note that

$$\begin{aligned} \mathsf{T}^+_{2k}(A) = \big|\{(a_1, \dots, a_{2k}, a'_1, \dots, a'_{2k}) \in A^{4k} : (a_1 + \dots + a_k) + (a_{k+1} + \dots + a_{2k}) \\ = (a'_1 + \dots + a'_k) + (a'_{k+1} + \dots + a'_{2k})\}\big| \\ = \sum_{x,y,z} r_{kA}(x) r_{kA}(y) r_{kA}(x+z) r_{kA}(y+z), \end{aligned} \tag{17}$$

so one can rewrite $\mathsf{T}^+_{2k}(A)$ via the additive energy of the function $r_{kA}(x)$. Secondly, for $k \ge 2$, we put

$$\mathsf{E}^+_k(A) = \sum_{x \in \mathbb{F}_p} (A \circ A)(x)^k = \sum_{x \in \mathbb{F}_p} r^k_{A-A}(x) = \mathsf{E}^+(\Delta_k(A), A^k), \tag{18}$$

where

$$\Delta_k(A) := \{(a, a, \dots, a) \in A^k\}.$$

Thus, $\mathsf{E}^+_2(A) = \mathsf{T}^+_2(A) = \mathsf{E}^+(A)$. Also, notice that we always have $|A|^k \le \mathsf{E}^+_k(A) \le |A|^{k+1}$ and moreover

$$\mathsf{E}^+_k(A) \le |A|^{k-l} \mathsf{E}^+_l(A) \quad \text{for all } l \le k. \tag{19}$$

Finally, let us remark that by definition (18) one has $\mathsf{E}^+_1(A) = |A|^2$. Some results about the properties of the energies E^+_k can be found in [Schoen and Shkredov 2013]. Sometimes we use $\mathsf{T}^+_k(f)$ and $\mathsf{E}^+_k(f)$ for an arbitrary function f and the first formula from (18) allows us to define $\mathsf{E}^+_k(A)$ for any positive k. It was proved in [Shkredov 2017, Proposition 16] that $(\mathsf{E}^+_k(f))^{1/2k}$ is a norm for even k and a real

function f. The fact that $(\mathsf{T}_k^+(f))^{1/2k}$ is a norm is contained in [Tao and Vu 2006] and follows from a generalization of inequality (15).

Let A be a set. Put

$$R[A] := \left\{ \frac{a_1 - a}{a_2 - a} : a, a_1, a_2 \in A, a_2 \neq a \right\}$$

and

$$Q[A] := \left\{ \frac{a_1 - a_2}{a_3 - a_4} : a_1, a_2, a_3, a_4 \in A, a_3 \neq a_4 \right\}.$$

All logarithms are base 2. The signs $\ll$ and $\gg$ are the usual Vinogradov symbols. When the constants in the signs depend on some parameter M, we write $\ll_M$ and $\gg_M$. For a positive integer n, we set $[n] = \{1, \dots, n\}$.

3. Preliminaries

We begin with a variation on the famous Plünnecke–Ruzsa inequality; see [Ruzsa 2009, Chapter 1].

Lemma 7. *Let $\boldsymbol{G}$ be a commutative group. Also, let $A, B_1, \dots, B_h \subseteq \boldsymbol{G}$, $|A + B_j| = \alpha_j |A|$, $j \in [h]$. Then there is a nonempty set $X \subseteq A$ such that*

$$|X + B_1 + \dots + B_h| \leq \alpha_1 \dots \alpha_h |X|. \tag{20}$$

Further for any $0 < \delta < 1$ there is $X \subseteq A$ such that $|X| \geq (1-\delta)|A|$ and

$$|X + B_1 + \dots + B_h| \leq \delta^{-h} \alpha_1 \dots \alpha_h |X|. \tag{21}$$

We need a result from [Rudnev 2017b] or see [Murphy et al. 2017, Theorem 8]. By the number of point-plane incidences $\mathcal{I}(\mathcal{P}, \Pi)$ between a set of points $\mathcal{P} \subseteq \mathbb{F}_p^3$ and a collection of planes Π in $\mathbb{F}_p^3$ we mean

$$\mathcal{I}(\mathcal{P}, \Pi) := \big|\{(p, \pi) \in \mathcal{P} \times \Pi : p \in \pi\}\big|.$$

Theorem 8. *Let p be an odd prime, $\mathcal{P} \subseteq \mathbb{F}_p^3$ be a set of points and Π be a collection of planes in $\mathbb{F}_p^3$. Suppose that $|\mathcal{P}| = |\Pi|$ and that k is the maximum number of collinear points in $\mathcal{P}$. Then the number of point-plane incidences satisfies*

$$\mathcal{I}(\mathcal{P}, \Pi) \ll \frac{|\mathcal{P}|^2}{p} + |\mathcal{P}|^{3/2} + k|\mathcal{P}|. \tag{22}$$

Notice that in $\mathbb{R}$ we do not need in the first term in estimate (22).

Let us derive a consequence of Theorem 8.

Lemma 9. *Let $A, Q \subseteq \mathbb{F}_p$ be two sets, $A, Q \neq \{0\}$, $M \geq 1$ be a real number, and $|QA| \leq M|Q|$. Then*

$$\mathsf{E}^+(Q) \leq C_* \left(\frac{M^2 |Q|^4}{p} + \frac{M^{3/2} |Q|^3}{|A|^{1/2}} \right), \tag{23}$$

where $C_* \geq 1$ is an absolute constant.

Proof. Put $A = A \setminus \{0\}$. We have

$$\begin{aligned}\mathsf{E}^+(Q) &= |\{q_1+q_2 = q_3+q_4 : q_1,q_2,q_3,q_4 \in Q\}| \\ &\le |A_*|^{-2}|\{q_1+\tilde{q}_2/a = q_3+\tilde{q}_4/a' : q_1,q_3 \in Q,\ \tilde{q}_2,\tilde{q}_4 \in QA,\ a,a' \in A_*\}|.\end{aligned}$$

The number of the solutions to the last equation can be interpreted as the number of incidences between the set of points $\mathcal{P} = Q \times QA \times A_*^{-1}$ and planes Π with $|\mathcal{P}| = |\Pi| = |A_*||Q||QA|$. Here $k = |QA|$ because $A, Q \ne \{0\}$. Using Theorem 8 and a trivial inequality $|QA| \le |Q||A|$, we obtain

$$\mathsf{E}^+(Q) \ll |A|^{-2}\Big(\frac{|A|^2|Q|^2|QA|^2}{p} + |Q|^{3/2}|QA|^{3/2}|A|^{3/2}\Big) \ll \frac{M^2|Q|^4}{p} + \frac{M^{3/2}|Q|^3}{|A|^{1/2}},$$

as required. □

Finally, we need a purely combinatorial Lemma 10. It is a new (for $k > 2$) and simple tool which allows us to estimate the restricted higher energy $\sum_{x\in P} r^k_{A-A}(x)$ via some energies of A and P; see (25), for example.

Lemma 10. *Let $\boldsymbol{G}$ be a finite abelian group and A, P subsets of $\boldsymbol{G}$. For any $k \ge 1$ one has*

$$\Big(\sum_{x\in P} r^k_{A-A}(x)\Big)^2 \le |A|^k \sum_x r^k_{A-A}(x) r_{P-P}(x). \tag{24}$$

In particular,

$$\Big(\sum_{x\in P} r^k_{A-A}(x)\Big)^4 \le |A|^{2k}\mathsf{E}^+_{2k}(A)\mathsf{E}^+(P). \tag{25}$$

Proof. Clearly, inequality (25) follows from (24) by the Cauchy–Schwarz inequality. To prove estimate (24), we observe that

$$\begin{aligned}\Big(\sum_{x\in P} r^k_{A-A}(x)\Big)^2 &= \Big(\sum_{x_1,\dots,x_k\in A} |P\cap(A-x_1)\cap\dots\cap(A-x_k)|\Big)^2 \\ &\le |A|^k \sum_{x_1,\dots,x_k} |P\cap(A-x_1)\cap\dots\cap(A-x_k)|^2 = |A|^k \sum_x r_{P-P}(x) r^k_{A-A}(x),\end{aligned}$$

as required. □

Combining Theorem 8 and Lemma 10, we obtain a corollary.

Corollary 11. *Let $A \subseteq \mathbb{F}_p$, and $B, P \subseteq \mathbb{F}_p^*$ be sets. Then for any $k \ge 1$ one has*

$$\Big(\sum_{x\in P} r^k_{A-A}(x)\Big)^4 \le C_*|A|^{2k}\mathsf{E}^+_{2k}(AB)\Big(\frac{|P|^4}{p} + \frac{|P|^3}{|B|^{1/2}}\Big). \tag{26}$$

Proof. By Lemma 10, we have

$$\Big(\sum_{x\in P} r^k_{A-A}(x)\Big)^2 \le |A|^k \sum_x r^k_{A-A}(x) r_{P-P}(x).$$

Further, clearly for any $b \in B$ we have

$$r_{A-A}(x) \le r_{AB-AB}(xb).$$

Hence

$$\Big(\sum_{x\in P} r^k_{A-A}(x)\Big)^2 \le \frac{|A|^k}{|B|}\sum_x\sum_{b\in B} r^k_{AB-AB}(xb) r_{P-P}(x) = \frac{|A|^k}{|B|}\sum_x r^k_{AB-AB}(x) r_{B(P-P)}(x).$$

Using the Cauchy–Schwarz inequality, we obtain

$$\Big(\sum_{x\in P} r^k_{A-A}(x)\Big)^4 \le \frac{|A|^{2k}}{|B|^2}\mathsf{E}^+_{2k}(AB)\sum_x r^2_{B(P-P)}(x).$$

To estimate the sum $\sum_x r^2_{B(P-P)}(x)$, we use Theorem 8 similar to the proof of Lemma 9 (see [Yazici et al. 2017]). Indeed, taking $\mathcal{P} = (p_1, b'p_2, b')$, $\Pi = (b, p'_1, bp_2)$, where $(b, b', p_1, p_2, p'_1, p'_2) \in B^2 \times P^4$, we have

$$\begin{aligned}\sum_x r^2_{B(P-P)}(x) &= |\{(b, b', p_1, p_2, p'_1, p'_2) \in B^2 \times P^4 : b(p_1 - p_2) = b'(p'_1 - p'_2)\}| \\ &= |\{(x, y, z) \in \mathcal{P},\ (b, p'_1, bp_2) \in \Pi : bx + y - p'_1 z = bp_2\}| = \mathcal{I}(\mathcal{P}, \Pi) \\ &\le C_*\Big(\frac{|B|^2|P|^4}{p} + |B|^{3/2}|P|^3\Big).\end{aligned}$$

Thus,

$$\Big(\sum_{x\in P} r^k_{A-A}(x)\Big)^4 \le C_*|A|^{2k}\mathsf{E}^+_{2k}(AB)\Big(\frac{|P|^4}{p} + \frac{|P|^3}{|B|^{1/2}}\Big). \qquad \square$$

4. Multiplicative subgroups

In this section we obtain the best upper bounds for $\mathsf{T}^+_k(\Gamma)$, $\mathsf{E}^+_k(\Gamma)$ and for the exponential sums over multiplicative subgroups Γ. We begin with the quantity $\mathsf{T}^+_k(\Gamma)$.

Theorem 12. *Let $\Gamma \subseteq \mathbb{F}_p^*$ be a multiplicative subgroup. Then for any $k \ge 2$, $2^{64k}C_*^4 \le |\Gamma|$ one has*

$$\mathsf{T}^+_{2^k}(\Gamma) \le 2^{4k+6}C_*\log^4|\Gamma| \cdot \frac{|\Gamma|^{2^{k+1}}}{p} + 16^{k^2}C_*^{k-1}\log^{4(k-1)}|\Gamma| \cdot |\Gamma|^{2^{k+1}-\frac{k+7}{2}}\mathsf{E}^+(\Gamma), \qquad (27)$$

where C_* is the absolute constant from Lemma 9.

Proof. Fix any $s \ge 2$. Our intermediate aim is to prove

$$\mathsf{T}^+_{2s}(\Gamma) \le 32C_*s^4\log^4|\Gamma| \cdot \Big(\frac{|\Gamma|^{4s}}{p} + |\Gamma|^{2s-1/2}\mathsf{T}^+_s(\Gamma)\Big). \qquad (28)$$

By (17), we have

$$\mathsf{T}^+_{2s}(\Gamma) = \sum_{x,y,z} r_{s\Gamma}(x) r_{s\Gamma}(y) r_{s\Gamma}(x+z) r_{s\Gamma}(y+z).$$

Put $\rho = \mathsf{T}^+_{2s}(\Gamma)/(16|\Gamma|^{3s})$. Since

$$\sum_{x,y,z:r_{s\Gamma}(x)\le\rho} r_{s\Gamma}(x)r_{s\Gamma}(y)r_{s\Gamma}(x+z)r_{s\Gamma}(y+z) \le \rho|\Gamma|^{3s} = \mathsf{T}^+_{2s}(\Gamma)/16$$

it follows that

$$\mathsf{T}^+_{2s}(\Gamma) \le \frac{4}{3}\sum\nolimits' k_{x,y,z} r_{s\Gamma}(x)r_{s\Gamma}(y)r_{s\Gamma}(x+z)r_{s\Gamma}(y+z) + \mathcal{E},$$

where the sum $\sum'$ is taken over nonzero variables x, y, z with $r_{s\Gamma}(x), r_{s\Gamma}(y), r_{s\Gamma}(x+z), r_{s\Gamma}(y+z) > \rho$ and

$$\mathcal{E} \le 4r_{s\Gamma}(0)\sum_{y,z} r_{s\Gamma}(y)r_{s\Gamma}(z)r_{s\Gamma}(y+z) \le 4r_{s\Gamma}(0)|\Gamma|^s\mathsf{T}^+_s(\Gamma) \le 4|\Gamma|^{2s-1}\mathsf{T}^+_s(\Gamma). \qquad (29)$$

Put $P_j = \{x : \rho 2^{j-1} < r_{s\Gamma}(x) \le \rho 2^j\} \subseteq \mathbb{F}_p^*$. If (28) does not hold, then, in particular, $\mathsf{T}^+_{2s}(\Gamma) \ge 2^5|\Gamma|^{2s-1/2}\mathsf{T}^+_s(\Gamma) \ge 2^5|\Gamma|^{3s-1/2}$ and hence the possible number of sets P_j does not exceed $L := s\log|\Gamma|$. Indeed, for any x one has $r_{s\Gamma}(x) \le |\Gamma|^{s-1}$ and hence $\rho 2^{j-1} = 2^{j-5}\mathsf{T}^+_{2s}(\Gamma)|\Gamma|^{-3s}$ must be less than $|\Gamma|^{s-1}$ otherwise the correspondent set P_j is empty. In other words,

$$2^{j-5} \le |\Gamma|^{4s-1}/\mathsf{T}^+_{2s}(\Gamma) \le |\Gamma|^{s-1/2}/2^5 \le |\Gamma|^s/2^5$$

as required. By the Dirichlet principle there is $\Delta = \rho 2^{j_0}$, and a set $P = P_{j_0}$ such that

$$\mathsf{T}^+_{2s}(\Gamma) \le \tfrac{4}{3}L^4(2\Delta)^4\mathsf{E}^+(P) + \mathcal{E} = \mathsf{T}'_{2s}(\Gamma) + \mathcal{E}.$$

Indeed, putting $f_i(x) = P_i(x)r_{s\Gamma}(x)$, and using (15), we get

$$\begin{aligned}\sum_{x,y,z}{}' r_{s\Gamma}(x)r_{s\Gamma}(y)r_{s\Gamma}(x+z)r_{s\Gamma}(y+z) &\le \sum_{i,j,k,l=1}^{L}\sum_{x,y,z} f_i(x)f_j(y)f_k(x+z)f_l(y+z)\\ &\le \sum_{i,j,k,l=1}^{L}(\mathsf{E}^+(f_i)\mathsf{E}^+(f_j)\mathsf{E}^+(f_k)\mathsf{E}^+(f_l))^{1/4}\\ &= \Big(\sum_{i=1}^{L}(\mathsf{E}^+(f_i))^{1/4}\Big)^4 \le L^3\sum_{i=1}^{L}\mathsf{E}^+(f_i) \le L^4\max_i \mathsf{E}^+(f_i).\end{aligned}$$

Moreover we always have $|P|\Delta^2 \le \mathsf{T}^+_s(\Gamma)$ and $|P|\Delta \le |\Gamma|^s$. Using Lemma 9, we obtain

$$\mathsf{E}^+(P) \le C_*\Big(\frac{|P|^4}{p} + \frac{|P|^3}{|\Gamma|^{1/2}}\Big).$$

Hence,

$$\mathsf{T}'_{2s}(\Gamma) \le \tfrac{4}{3}(16C_*)L^4\Delta^4\Big(\frac{|P|^4}{p} + \frac{|P|^3}{|\Gamma|^{1/2}}\Big) \le \tfrac{4}{3}(16C_*)L^4\Big(\frac{|\Gamma|^{4s}}{p} + \frac{|P|^3\Delta^4}{|\Gamma|^{1/2}}\Big). \qquad (30)$$

Let us consider the second term in (30). Then in view of $|P|\Delta^2 \le \mathsf{T}^+_s(\Gamma)$ and $|P|\Delta \le |\Gamma|^s$, we have

$$|P|^3\Delta^4 = (P\Delta)^2 P\Delta^2 \le |\Gamma|^{2s}\mathsf{T}^+_s(\Gamma).$$

In other words, by (29), we get

$$\mathsf{T}^+_{2s}(\Gamma) \le \tfrac{4}{3}(16C_*)L^4\left(\frac{|\Gamma|^{4s}}{p}+|\Gamma|^{2s-1/2}\mathsf{T}^+_s(\Gamma)\right)+4|\Gamma|^{2s-1}\mathsf{T}^+_s(\Gamma)$$
$$\le 32C_*s^4\log^4|\Gamma|\cdot\left(\frac{|\Gamma|^{4s}}{p}+|\Gamma|^{2s-1/2}\mathsf{T}^+_s(\Gamma)\right)$$

and inequality (28) is proved.

Now applying formula (28) successively $k-1$ times, we obtain

$$\mathsf{T}^+_{2^k}(\Gamma) \le 2^{4k+6}C_*\log^4|\Gamma|\cdot\frac{|\Gamma|^{2^{k+1}}}{p}+16^{k^2}C_*^{k-1}\log^{4(k-1)}|\Gamma|\cdot|\Gamma|^{2^k+\cdots+4-\frac{k-1}{2}}\mathsf{E}^+(\Gamma)$$
$$\le 2^{4k+6}C_*\log^4|\Gamma|\cdot\frac{|\Gamma|^{2^{k+1}}}{p}+16^{k^2}C_*^{k-1}\log^{4(k-1)}|\Gamma|\cdot|\Gamma|^{2^{k+1}-\frac{k+7}{2}}\mathsf{E}^+(\Gamma). \qquad (31)$$

To get the first term in the last formula we have used our condition $2^{64k}C_*^4 \le |\Gamma|$ to ensure that $|\Gamma|^{1/2} \ge 2^{4k+1}C_*\log^4|\Gamma|$. □

Remark 13. The condition $2^{64k}C_*^4 \le |\Gamma|$ can be dropped, but in that case we will have the factor $16^{k^2}(C_*\log|\Gamma|)^{k-1}$ in the first term of (27).

Splitting any Γ-invariant set onto cosets over Γ and applying the norm property of T^+_l, we obtain:

Corollary 14. *Let $\Gamma \subseteq \mathbb{F}_p^*$ be a multiplicative subgroup, and $Q \subseteq \mathbb{F}_p^*$ be a set with $Q\Gamma = Q$. Then for any $k \ge 2$, $2^{64k}C_*^4 \le |\Gamma|$ one has*

$$\mathsf{T}^+_{2^k}(Q) \le 2^{4k+6}C_*\log^4|\Gamma|\cdot\frac{|Q|^{2^{k+1}}}{p}+16^{k^2}C_*^{k-1}\log^{4(k-1)}|\Gamma|\cdot|\Gamma|^{-\frac{k+7}{2}}\mathsf{E}^+(\Gamma)|Q|^{2^{k+1}}. \qquad (32)$$

Let Γ be a subgroup of size less than $\sqrt{p}$. Considering the particular case $k=2$ of the formula in Theorem 12 and using $\mathsf{E}^+(\Gamma) \ll |\Gamma|^{5/2-c}$, where $c>0$ is an absolute constant (see [Shkredov 2013]), one has:

Corollary 15. *Let Γ be a multiplicative subgroup, $|\Gamma| \le \sqrt{p}$. Then*

$$\mathsf{T}^+_4(\Gamma) \ll \frac{|\Gamma|^8\log^4|\Gamma|}{p}+|\Gamma|^{6-c}.$$

In particular, $|4\Gamma| \gg |\Gamma|^{2+c}$.

Previous results on $\mathsf{T}^+_k(\Gamma)$, $|\Gamma| \le \sqrt{p}$ with small k had the form $\mathsf{T}^+_k(\Gamma) \ll |\Gamma|^{2k-2+c_k}$ with some $c_k>0$; see, e.g., [Konyagin and Shparlinski 1999]. The best upper bound for $\mathsf{T}^+_3(\Gamma)$ can be found in [Shteinikov 2015].

Now we prove a corollary about exponential sums over subgroups, which is parallel to results from [Bourgain and Garaev 2009; Bourgain et al. 2006; Garaev 2010]. The difference between the previous estimates and Corollary 16 is just a slightly better constant C in (34).

Corollary 16. *Let Γ be a multiplicative subgroup, $|\Gamma| \ge p^{\delta}$, $\delta > 0$. Then for all sufficiently large p one has*

$$\max_{\xi \ne 0} |\hat{\Gamma}(\xi)| \ll |\Gamma| \cdot p^{-\delta/2^{7+2\delta^{-1}}}. \tag{33}$$

Further we have a nontrivial upper bound $o(|\Gamma|)$ for the maximum in (33) *if*

$$\log |\Gamma| \ge \frac{C \log p}{\log \log p}, \tag{34}$$

where $C > 2$ is any constant.

Proof. We can assume that $|\Gamma| < \sqrt{p}$, say, because otherwise the estimate (33) is known; see [Konyagin and Shparlinski 1999]. By ρ denote the maximum in (33). Then by Theorem 12, a trivial bound $\mathsf{E}^+(\Gamma) \le |\Gamma|^3$ and (16), we obtain

$$|\Gamma|\rho^{2^{k+1}} \le p\mathsf{T}_{2^k}(\Gamma) \le 2^{4k+6} C_* \log^4 |\Gamma| \cdot |\Gamma|^{2^{k+1}} + 16^{k^2} C_*^{k-1} \log^{4(k-1)} |\Gamma| \cdot |\Gamma|^{2^{k+1}-(k+1)/2} p, \tag{35}$$

provided $2^{64k} C_*^4 \le |\Gamma|$. Put $k = \lceil 2 \log p / \log |\Gamma| + 4 \rceil \le 2/\delta + 5$. Also, notice that

$$\frac{p \log^{4(k-1)} |\Gamma|}{|\Gamma|^{k/2}} \le 1, \tag{36}$$

because $k \ge 2 \log p / \log |\Gamma| + 4$ and p is a sufficiently large number depending on δ (the choice of k is slightly larger than $2 \log p / \log |\Gamma|$ to "kill" p by division by $|\Gamma|^{k/2}$ as well as logarithms $\log^{4(k-1)} |\Gamma|$). Also, since $|\Gamma| \ge p^{\delta}$, it follows that $2^{64k} C_*^4 \le |\Gamma|$ for sufficiently large p. Taking a power $1/2^{k+1}$ from both parts of (35), we see in view of (36) that

$$\rho \ll |\Gamma|(|\Gamma|^{-1/2^{k+2}} + |\Gamma|^{-1/2^{k+2}}) \ll |\Gamma|^{1-1/2^{k+2}} \ll |\Gamma| \cdot p^{-\frac{\delta}{2^{7+2\delta^{-1}}}}.$$

To prove the second part of our corollary just notice that the same choice of k gives something nontrivial if $2^{k+2} \le \varepsilon \log |\Gamma|$ for any $\varepsilon > 0$. In other words, it is enough to have

$$k + 2 \le \frac{2 \log p}{\log |\Gamma|} + 7 \le \log \log |\Gamma| - \log(1/\varepsilon).$$

It means that the inequality $\log |\Gamma| \ge C \log p / (\log \log p)$ for any $C > 2$ is enough. □

Remark 17. One can improve some constants in the proof (but not the constant C in (34)), probably, but we did not make such calculations.

Now we estimate a "dual" quantity $\mathsf{E}_s^+(Q)$ for Γ-invariant set Q (about duality of $\mathsf{T}_{k/2}^+(A)$ and $\mathsf{E}_k^+(A)$; see [Schoen and Shkredov 2013] and (40)–(43)). We give even two bounds and both of them use the Fourier approach. Our first estimate (37) relatively quickly follows from Corollary 14 and the price for it is the appearance p in the bounds. The second estimate (39) is more delicate but requires more work.

Theorem 18. *Let $\Gamma \subseteq \mathbb{F}_p^*$ be a multiplicative subgroup, and $Q \subseteq \mathbb{F}_p^*$ be a set with $Q\Gamma = Q$ and $|Q|^2|\Gamma| \le p^2$. Then for $0 \le k$, $2^{64k} C_*^4 \le |\Gamma|$ one has*

$$\mathsf{E}^+_{2^{k+1}}(Q) \le 2^{2^{k+2}+3} (\log |Q|)^{2^{k+1}} |Q|^{2^{k+1}} \big(2^{4k+6} C_* \log^4 |Q| + 16^{k^2} C_*^{k-1} \log^{4(k-1)} |Q| \cdot |\Gamma|^{-\frac{k+1}{2}} p\big). \tag{37}$$

Further let $k \ge 1$ be such that

$$\Gamma|^{(k+2)/2} \ge |Q| \log^{4k} |Q|. \quad (38)$$

Then

$$\mathsf{E}^+_{2^{k+1}}(Q) \le (2^8 C_*)^{k+1} |Q|^{2^{k+1}} |\Gamma|^{1/2}. \quad (39)$$

Proof. We begin with (37) and we prove this inequality by induction. For $k = 0$ the result is trivial in view of our condition $|Q|^2|\Gamma| \le p^2$. Put $s = 2^k$, $k \ge 1$. By the Parseval identity and (12), we have

$$\mathsf{E}^+_{2s}(Q) = \frac{1}{p^{2s-1}} \sum_{x_1+\dots+x_{2s}=0} |\hat{Q}(x_1)|^2 \dots |\hat{Q}(x_{2s})|^2 \quad (40)$$

$$\le \frac{2s|Q|^2 \mathsf{E}^+_{2s-1}(Q)}{p} + \frac{1}{p^{2s-1}} \sum_{\substack{x_1+\dots+x_{2s}=0 \\ x_j \ne 0 \text{ for all } j}} |\hat{Q}(x_1)|^2 \dots |\hat{Q}(x_{2s})|^2 \quad (41)$$

$$= \frac{2s|Q|^2 \mathsf{E}^+_{2s-1}(Q)}{p} + \mathsf{E}'_{2s}(Q). \quad (42)$$

Put $L = \log |Q|$. By the Parseval identity

$$\frac{1}{p^{2s-1}} \sum_{\substack{x_1+\dots+x_{2s}=0 \\ x_j \ne 0 \text{ for all } j}} |\hat{Q}(x_1)|^2 \dots |\hat{Q}(x_{2s})|^2$$

$$\le \max_{x \ne 0} |\hat{Q}(x)|^2 \cdot \frac{1}{p^{2s-1}} \sum_{\substack{x_1+\dots+x_{2s}=0 \\ x_j \ne 0 \text{ for all } j}} |\hat{Q}(x_1)|^2 \dots |\hat{Q}(x_{2s-1})|^2 \le \max_{x \ne 0} |\hat{Q}(x)|^2 \cdot |Q|^{2s-1}.$$

Hence, as in the proof of Theorem 12, consider $\rho^2 = \mathsf{E}^+_{2s}(Q)/(4s|Q|^{2s-1})$ and the sets

$$P_j = \{x : \rho 2^{j-1} < |\hat{Q}(x)| \le \rho 2^j\} \subseteq \mathbb{F}_p^*.$$

Using the Dirichlet principle, we find $\Delta = \rho 2^{j_0} \ge \rho$ and $P = P_{j_0}$ such that

$$\mathsf{E}'_{2s}(Q) \le \frac{4L^{2s}(2\Delta)^{4s}}{p^{2s-1}} \mathsf{T}^+_s(P). \quad (43)$$

Here we bound the number of sets P_j by the number L because of

$$2^{2j-2} \le |Q|^2/\rho^2 \le 4s|Q|^{2s+1}/\mathsf{E}^+_{2s}(Q) \le |Q|/4$$

and the last inequality follows if (37) does not hold. Clearly, $P\Gamma = P$ (and this is the crucial point of the proof, actually). Applying Corollary 14, we get

$$\mathsf{E}'_{2s}(Q) \le$$

$$\frac{2^{4s+2}L^{2s}\Delta^{4s}}{p^{2s-1}} \left(2^{4k+6} C_* \log^4 |\Gamma| \cdot \frac{|P|^{2s}}{p} + 16^{k^2} C_*^{k-1} \log^{4(k-1)} |\Gamma| \cdot |\Gamma|^{-\frac{k+7}{2}} \mathsf{E}^+(\Gamma) |P|^{2s} \right). \quad (44)$$

By the Parseval identity, we see that

$$\Delta^2 |P| \le |Q| p. \quad (45)$$

Hence

$$\mathsf{E}'_{2s}(Q) \le 2^{4s+2}L^{2s}|Q|^{2s}\cdot(2^{4k+6}C_*L^4+16^{k^2}C_*^{k-1}L^{4(k-1)}\cdot|\Gamma|^{-\frac{k+7}{2}}\mathsf{E}^+(\Gamma)p). \qquad (46)$$

Using a trivial bound $\mathsf{E}^+(\Gamma)\le|\Gamma|^3$, we get

$$\mathsf{E}'_{2s}(Q) \le 2^{4s+2}L^{2s}|Q|^{2s}\cdot(2^{4k+6}C_*L^4+16^{k^2}C_*^{k-1}L^{4(k-1)}\cdot|\Gamma|^{-\frac{k+1}{2}}p). \qquad (47)$$

Applying a crude bound (19), namely, $\mathsf{E}^+_{2s-1}(Q)\le|Q|^{s-1}\mathsf{E}^+_s(Q)$, the condition $|Q|^2|\Gamma|\le p^2$, and induction assumption, we get

$$\begin{aligned}\frac{2s|Q|^2\mathsf{E}^+_{2s-1}(Q)}{p} &\le \frac{2s|Q|^{s+1}\mathsf{E}^+_s(Q)}{p}\\ &\le \frac{2s|Q|^{s+1}}{p}\cdot L^s|Q|^s\cdot 2^{2s+3}(2^{4k+2}C_*L^4+16^{(k-1)^2}C_*^{k-2}L^{4(k-2)}\cdot|\Gamma|^{-k/2}p)\\ &\le 2^{4s+2}L^{2s}|Q|^{2s}\cdot(2^{4k+6}C_*L^4+16^{k^2}C_*^{k-1}L^{4(k-1)}\cdot|\Gamma|^{-\frac{k+1}{2}}p).\end{aligned}$$

Hence combining the last estimate with (47), we derive

$$\mathsf{E}^+_{2^{k+1}}(Q)\le 2^{2^{k+2}+3}L^{2^{k+1}}|Q|^{2^{k+1}}\cdot(2^{4k+6}C_*L^4+16^{k^2}C_*^{k-1}L^{4(k-1)}\cdot|\Gamma|^{-\frac{k+1}{2}}p)$$

and thus we have obtained (37).

To get (39), put $l=2^{k-1}$, $k\ge 1$ and consider $\mathsf{E}^+_{4l}(Q)$. Further define $g(x)=r^l_{Q-Q}(x)$ and notice that $\hat{g}(\xi)\ge 0$, $\hat{g}(0)=\mathsf{E}^+_l(Q)$. Moreover, taking the Fourier transform as in (40) and using the Dirichlet principle, we get

$$\begin{aligned}\mathsf{E}^+_{4l}(Q)=\sum_x(Q\circ Q)^{4l}(x)=\sum_x g^4(x)=\frac{\mathsf{E}^+(\hat{g})}{p^3}=\frac{1}{p^3}\sum_{x,y,z}\hat{g}(x)\hat{g}(y)\hat{g}(x+z)\hat{g}(y+z)\\ \le\frac{4\hat{g}(0)}{p^3}\sum_{y,z}\hat{g}(y)\hat{g}(z)\hat{g}(y+z)+\frac{1}{p^3}\sum_{x\ne 0,y\ne 0,z\ne 0}\hat{g}(x)\hat{g}(y)\hat{g}(x+z)\hat{g}(y+z)\\ \le\frac{4\mathsf{E}^+_l(Q)\mathsf{E}^+_{3l}(Q)}{p}+\frac{4L^4(2\omega)^4}{p^3}\mathsf{E}^+(G), \qquad (48)\end{aligned}$$

where $G=\{\xi:\omega<\hat{g}(\xi)\le 2\omega\}\subseteq\mathbb{F}_p^*$, and $\omega\ge 2^{-3}\mathsf{E}^+_{4l}(Q)|Q|^{-3l}:=\rho_*$ because the sum over $\hat{g}(\xi)<\rho_*$ by (10) does not exceed

$$\frac{4\rho_*}{p^3}\cdot\sum_{x,y,z}\hat{g}(y)\hat{g}(x+z)\hat{g}(y+z)=4\rho_*g^3(0)=4\rho_*|Q|^{3l}.$$

Further in view of the Parseval identity, we see that

$$\omega^2|G|\le\sum_{\xi\in G}\hat{g}(\xi)^2\le p\mathsf{E}^+_{2l}(Q), \qquad (49)$$

and by (10),

$$\omega|G|\le\sum_{\xi\in G}\hat{g}(\xi)=pg(0)=p|Q|^l. \qquad (50)$$

Clearly, G is a Γ-invariant set (again, this is the crucial point of the proof). Further returning to (48) and applying Lemma 9, we see that

$$\mathsf{E}^+_{4l}(Q) \le \frac{4\mathsf{E}^+_l(Q)\mathsf{E}^+_{3l}(Q)}{p} + \frac{2^6L^4\omega^4}{p^3}\mathsf{E}^+(G) \le \frac{4\mathsf{E}^+_l(Q)\mathsf{E}^+_{3l}(Q)}{p} + \frac{2^6C_*L^4\omega^4}{p^3}\Bigl(\frac{|G|^4}{p} + \frac{|G|^3}{|\Gamma|^{1/2}}\Bigr)$$
$$= \frac{4\mathsf{E}^+_l(Q)\mathsf{E}^+_{3l}(Q)}{p} + \mathsf{E}'_{4l}(Q).$$

Applying (49) and (50), we get

$$\mathsf{E}'_{4l}(Q) \le 2^6C_*L^4|Q|^{4l} + \frac{2^6C_*L^4(\omega|G|)^2\omega^2|G|}{|\Gamma|^{1/2}p^3} \le 2^6C_*L^4|Q|^{4l} + 2^6C_*L^4|Q|^{2l}\mathsf{E}^+_{2l}(Q)|\Gamma|^{-1/2}.$$

It follows that

$$\mathsf{E}^+_{4l}(Q) \le \frac{4\mathsf{E}^+_l(Q)\mathsf{E}^+_{3l}(Q)}{p} + 2^6C_*L^4|Q|^{2l}\mathsf{E}^+_{2l}(Q)\Bigl(\frac{|Q|^{2l}}{\mathsf{E}^+_{2l}(Q)} + \frac{1}{|\Gamma|^{1/2}}\Bigr). \qquad (51)$$

Further estimating the first term of (51) very roughly as

$$\frac{\mathsf{E}^+_l(Q)\mathsf{E}^+_{3l}(Q)}{p} \le \frac{|Q|^{l+1}\mathsf{E}^+_{3l}(Q)}{p} \le \frac{|Q|^{2l+1}\mathsf{E}^+_{2l}(Q)}{p},$$

we get in view of our condition $|Q|^2|\Gamma| \le p^2$ that this term is less than $L^4|Q|^{2l}\mathsf{E}^+_{2l}(Q)|\Gamma|^{-1/2}$. Hence

$$\mathsf{E}^+_{4l}(Q) \le 2^7C_*L^4|Q|^{2l}\mathsf{E}^+_{2l}(Q)\Bigl(\frac{|Q|^{2l}}{\mathsf{E}^+_{2l}(Q)} + \frac{1}{|\Gamma|^{1/2}}\Bigr). \qquad (52)$$

Notice that the term $|Q|^{2l}/\mathsf{E}^+_{2l}(Q) + 1/|\Gamma|^{1/2} \le 2\cdot\max\{|Q|^{2l}/\mathsf{E}^+_{2l}(Q), 1/|\Gamma|^{1/2}\} \le 2$. Applying bound (52) exactly $0 \le s \le k$ times, where s is the maximal number (if it exists) such that the second term $1/|\Gamma|^{1/2}$ in (52) dominates, we obtain

$$\mathsf{E}^+_{2^{k+1}}(Q) \le (2^8C_*)^sL^{4s}|\Gamma|^{-s/2}|Q|^{2^k+\ldots+2^{k-s+1}}\mathsf{E}^+_{2^{k-s+1}}(Q)\Bigl(\frac{|Q|^{2^{k-s+1}}}{\mathsf{E}^+_{2^{k-s+1}}(Q)} + \frac{1}{|\Gamma|^{1/2}}\Bigr). \qquad (53)$$

Now by the definition of s, we see that the first term in (53) dominates. Hence, using (51), (52) one more time (if $s < k$), we get

$$\mathsf{E}^+_{2^{k+1}}(Q) \le 2(2^8C_*)^sL^{4s}|\Gamma|^{-s/2}|Q|^{2^{k+1}-2^{k-s+1}}\cdot|Q|^{2^{k-s+1}} = 2(2^8C_*)^sL^{4s}|\Gamma|^{-s/2}|Q|^{2^{k+1}}. \qquad (54)$$

From the assumption $|\Gamma|^{(k+2)/2} \ge |Q|\log^{4k}|Q|$, it follows that $|\Gamma| \ge |Q|^{2/(k+2)}\log^{8k/(k+2)}|Q|$. Hence bound (54) is much better than (39) if $s < k$. If $s = k$, then by the same calculations, we derive

$$\mathsf{E}^+_{2^{k+1}}(Q) \le (2^8C_*)^kL^{4k}|\Gamma|^{-k/2}\mathsf{E}^+_2(Q)|Q|^{2^{k+1}-2}.$$

Since $|Q|^2|\Gamma| \le p^2$ by Lemma 9, it follows that $\mathsf{E}^+(Q) \le 2C_*|Q|^3/|\Gamma|^{1/2}$ and hence

$$\mathsf{E}^+_{2^{k+1}}(Q) \le (2^8C_*)^{k+1}L^{4k}|\Gamma|^{-(k+1)/2}|Q|^{2^{k+1}+1}.$$

Further by the choice of k, namely, $|\Gamma|^{(k+2)/2} \ge |Q|\log^{4k}|Q|$ we see that the last bound is better than (39). Finally, if $s=0$, then by definition $\mathsf{E}^+_{2^k}(Q) \le |Q|^{2^k}|\Gamma|^{1/2}$ and hence $\mathsf{E}^+_{2^{k+1}}(Q) \le |Q|^{2^{k+1}}|\Gamma|^{1/2}$. $\square$

Remark 19. From the second part of the arguments above one can derive explicit bounds for the energies $\mathsf{E}^+_s(Q)$ for small s. For example,

$$\mathsf{E}_4(Q) \ll \frac{|Q|^2\mathsf{E}_3(Q)}{p} + (\log|\Gamma|)^4|Q|^4 + (\log|\Gamma|)^4|Q|^2\mathsf{E}(Q)|\Gamma|^{-1/2}.$$

Now we obtain a uniform upper bound for the size of the intersection of an additive shift of any Γ-invariant set. Our bound (56) is especially effective if the sizes of Q_1, Q_2 are comparable with the size of Γ, namely, $|Q_1|, |Q_2| \ll |\Gamma|^C$, where C is an absolute constant (which can be large). In this case the number k below is a constant as well.

Corollary 20. *Let $\Gamma \subseteq \mathbb{F}_p^*$ be a multiplicative subgroup, $|\Gamma| \ge p^\delta$, $\delta > 0$, and $Q_1, Q_2 \subseteq \mathbb{F}_p^*$ be two sets with $Q_1\Gamma = Q_1$, $Q_2\Gamma = Q_2$, $|Q_1|^2|\Gamma| \le p^2$, $|Q_2|^2|\Gamma| \le p^2$. Put $Q = \max\{|Q_1|, |Q_2|\}$. Then for any $x \ne 0$, one has*

$$|Q_1 \cap (Q_2 + x)| \ll \sqrt{|Q_1||Q_2|}\log Q \cdot p^{-\delta/2^{7+2\delta^{-1}}}. \qquad (55)$$

Further choose $k \ge 1$ such that $|\Gamma|^{(k+2)/2} \ge Q\log^{4k} Q$. Then, for an arbitrary $x \ne 0$,

$$|Q_1 \cap (Q_2 + x)| \ll \sqrt{|Q_1||Q_2|}\cdot|\Gamma|^{-1/4\cdot 2^{-k}}. \qquad (56)$$

Proof. From the conditions $|Q_1|^2|\Gamma| \le p^2$, $|Q_2|^2|\Gamma| \le p^2$, it follows that $|\Gamma| \le p^{2/3}$. Put $L = \log Q$. On the one hand, applying the Cauchy–Schwarz inequality, we obtain

$$\sum_y r^{2^{k+1}}_{Q_1-Q_2}(y) \le (\mathsf{E}^+_{2^{k+1}}(Q_1))^{1/2}(\mathsf{E}^+_{2^{k+1}}(Q_2))^{1/2}.$$

On the other hand, by formula (37) of Theorem 18 and Γ-invariance of Q_1, Q_2, we have

$$\begin{aligned}|\Gamma||Q_1 \cap (Q_2+x)|^{2^{k+1}} &\le \sum_y r^{2^{k+1}}_{Q_1-Q_2}(y)\\ &\le 2^{2^{k+2}+3}L^{2^{k+1}}(|Q_1||Q_2|)^{2^k}(2^{4k+6}L^4 + 16^{k^2}C_*^{k-1}L^{4(k-1)}\cdot|\Gamma|^{-\frac{k+1}{2}}p),\end{aligned}$$

provided $2^{64k}C_*^4 \le |\Gamma|$. As in Corollary 16 choosing $k = \lceil 2\log p/\log|\Gamma| + 4\rceil \le 2/\delta + 5$ and applying an analogue of (36) which holds for large p, namely,

$$\frac{pL^{4(k-1)}}{|\Gamma|^{k/2}} \ll 1$$

we obtain

$$\begin{aligned}|Q_1 \cap (Q_2+x)| &\ll L\sqrt{|Q_1||Q_2|}\cdot(|\Gamma|^{-1/2^{k+2}} + |\Gamma|^{-1/2^{k+2}})\\ &\ll L\sqrt{|Q_1||Q_2|}|\Gamma|^{-1/2^{k+2}} \ll L\sqrt{|Q_1||Q_2|}p^{-\delta/2^{7+2\delta^{-1}}},\end{aligned}$$

and it easy to ensure that inequality $2^{64k}C_*^4 \le |\Gamma|$ takes place for sufficiently large p.

To derive (56), we just use the second formula (39) of Theorem 18 and the previous calculations. $\square$

Remark 21. It is known (see, e.g., [Konyagin and Shparlinski 1999]) that if $\Gamma \subseteq \mathbb{F}_p^*$ is a multiplicative subgroup with $|\Gamma| < p^{3/4}$, then for any $x \neq 0$ one has $|\Gamma \cap (\Gamma + x)| \ll |\Gamma|^{2/3}$ and this bound is tight in some regimes. One can extend this to larger Γ-invariant sets and obtain a lower bound of a comparable quality. It gives a lower estimate in (55).

Indeed, let $\Gamma \subseteq \mathbb{F}_p^*$ be a multiplicative subgroup with $|\Gamma| < p^{1/2}$. Consider $R = R[\Gamma]$ and $Q = Q[\Gamma]$. It was proved in [Shkredov 2016b] that $|R| \gg |\Gamma|^2/\log|\Gamma|$ and one can check that $R = 1 - R$; see, e.g., [Murphy et al. 2017]. Finally, the set Q is Γ-invariant and it is easy to check [Shkredov 2016a] that $|Q| \leq |\Gamma|^3$. Hence

$$|Q \cap (1-Q)| \geq |R| \gg \frac{|\Gamma|^2}{\log|\Gamma|} \gg \frac{|Q|^{2/3}}{\log|Q|}.$$

Also, notice that if $|\Gamma| < p^{1/2}$ and $|Q[\Gamma]|^2|\Gamma| \leq p^2$, then $|Q[\Gamma]| \gg |\Gamma|^{2+c}$ for some $c > 0$; see the first part of Corollary 35 from the next section.

Corollary 20 gives a nontrivial upper bound for the common additive energy of an arbitrary invariant set and *any* subset of $\mathbb{F}_p$.

Corollary 22. *Let $\Gamma \subseteq \mathbb{F}_p^*$ be a multiplicative subgroup, $|\Gamma| \geq p^\delta$, $\delta > 0$, and $Q \subseteq \mathbb{F}_p^*$ be a set with $Q\Gamma = Q$, $|Q|^2|\Gamma| \leq p^2$. Then for any set $A \subseteq \mathbb{F}_p$, one has*

$$\mathsf{E}^+(A,Q) \ll |Q||A|^2 \cdot p^{-\delta/2^{7+2\delta^{-1}}} \log|Q| + |A||Q|. \tag{57}$$

Further, for an arbitrary $\alpha \neq 0$,

$$\mathsf{E}^\times(A,Q+\alpha) \ll |Q||A|^2 \cdot p^{-\delta/2^{7+2\delta^{-1}}} \log|Q| + |A||Q|. \tag{58}$$

In particular,

$$|A+Q| \gg |Q| \cdot \min\{|A|, p^{\frac{\delta}{2^{7+2\delta^{-1}}}} \log^{-1}|Q|\}, \tag{59}$$

and

$$|A(Q+\alpha)| \gg |Q| \cdot \min\{|A|, p^{\frac{\delta}{2^{7+2\delta^{-1}}}} \log^{-1}|Q|\}. \tag{60}$$

If $k \geq 1$ is chosen as $|\Gamma|^{(k+2)/2} \geq |Q|\log^{4k}|Q|$, then one can replace the quantity $p^{\frac{\delta}{2^{7+2\delta^{-1}}}} \log^{-1}|Q|$ above by $|\Gamma|^{-1/4\cdot 2^{-k}}$.

Proof. Inequalities (59), (60) follow from (57), (58) via the Cauchy–Schwarz inequality, so it is enough to obtain the required upper bound for the additive energy of A and Q and for the multiplicative energy of A and $Q+\alpha$. By Corollary 20, we have

$$\mathsf{E}^+(A,Q) = \sum_x r_{A-A}(x) r_{Q-Q}(x) = |A||Q| + \sum_{x\neq 0} r_{A-A}(x) r_{Q-Q}(x)$$
$$\ll |A||Q| + |Q||A|^2 \cdot p^{-\delta/2^{7+2\delta^{-1}}} \log|Q|,$$

as required. Similarly

$$\mathsf{E}^\times(A,Q+\alpha) \ll |A||Q| + \sum_{x\neq 0,1} r_{A/A}(x) r_{(Q+\alpha)/(Q+\alpha)}(x) \ll |A||Q| + |Q||A|^2 \cdot p^{-\delta/2^{7+2\delta^{-1}}} \log|Q|,$$

because in view of Corollary 20 one has

$$r_{(Q+\alpha)/(Q+\alpha)}(x) = |Q \cap (xQ + \alpha(x-1))| \ll |Q| \cdot p^{-\delta/2^{7+2\delta^{-1}}} \log |Q|.$$

So, we have obtained bounds (57)–(60) with $p^{\frac{\delta}{2^{7+2\delta^{-1}}}} \log^{-1} |Q|$, and to replace it by $|\Gamma|^{-1/4\cdot 2^{-k}}$ one should use the second part of Corollary 20. □

From (59) one can obtain that for any multiplicative subgroup $\Gamma \subseteq \mathbb{F}_p^*$ there is N such that $N\Gamma = \mathbb{F}_p$ and $N \ll \delta^{-1}4^{\delta^{-1}}$. The results of comparable quality were obtained in [Glibichuk and Konyagin 2007].

5. The proof of the main result

In this section we obtain an upper bound for $\mathsf{T}_k^+(A)$ (see Theorem 23) and an upper bound for $\mathsf{E}_k^+(A)$ (see Theorem 27) in the case when the size of the product set AB is small comparable to A, where B is a sufficiently large set. From the last result we derive our quantitative asymmetric sum-product Theorem 5 from the introduction. Let us begin with an upper bound for $\mathsf{T}_k^+(A)$.

Theorem 23. *Let $A, B \subseteq \mathbb{F}_p$ be sets, $M \geq 1$ be a real number, and $|AB| \leq M|A|$, $|A| > 1$. Then for any $k \geq 2$, $2^{16k} M^{2^{k+1}} C_*^2 \log^8 |A| \leq |B|$, one has*

$$\mathsf{T}_{2^k}^+(A) \leq 2^{4k+6} C_* \log^4 |A| \cdot \frac{M^{2^k} |A|^{2^{k+1}}}{p} + 16^{k^2} C_*^{k-1} M^{2^{k+1}} \log^{4(k-1)} |A| \cdot |A|^{2^{k+1}-4} |B|^{-\frac{k-1}{2}} \mathsf{E}^+(A). \tag{61}$$

Proof. We have $B \neq \{0\}$ by the condition $2^{16k} M^{2^{k+1}} C_*^2 \log^8 |A| \leq |B|$, for instance. We apply the arguments and the notation of the proof of Theorem 12. Fix any $s \geq 2$ and put $L := s \log |A|$. Our intermediate aim is to prove

$$\mathsf{T}_{2s}^+(A) \leq C s^4 M^{2s} \log^4 |A| \cdot \left(\frac{|A|^{4s}}{p} + \frac{|A|^{2s}}{\sqrt{|B|}} \mathsf{T}_s^+(A) \right), \tag{62}$$

where $C = 2^5 C_*$. As in the proof of Theorem 12, we get

$$\mathsf{T}_{2s}^+(A) \leq \tfrac{4}{3} L^4 (2\Delta)^4 \mathsf{E}^+(P) + \mathcal{E},$$

where

$$\mathcal{E} \leq 4|A|^{2s-1} \mathsf{T}_s^+(A). \tag{63}$$

Further, $\Delta > \mathsf{T}_{2s}^+(A)/(16|A|^{3s})$ is a real number and $P = \{x : \Delta < r_{sA}(x) \leq 2\Delta\} \subseteq \mathbb{F}_p^*$. Moreover, we always have $|P|\Delta^2 \leq \mathsf{T}_s^+(A)$. Notice also

$$|P|\Delta \leq \sum_{x \in P} r_{sA}(x) \leq \sum_x r_{sA}(x) \leq |A|^s.$$

To proceed as in the proof of Theorem 12, we need to estimate $|PB|$. Observe that for any $x \in PB$ one has $r_{sAB}(x) \geq \Delta$. Thus, we have

$$|PB|\Delta \leq \sum_{x \in PB} r_{sAB}(x) \leq |AB|^s \leq M^s |A|^s. \tag{64}$$

Hence using Lemma 9, we obtain

$$\mathsf{E}^+(P) \le C_*\Big(\frac{M^{2s}|A|^{2s}|P|^2}{\Delta^2 p} + \frac{M^{3s/2}|A|^{3s/2}|P|^{3/2}}{\Delta^{3/2}|B|^{1/2}}\Big).$$

Hence in view of estimate (63), combining with $|P|\Delta \le |A|^s$ and $|P|\Delta^2 \le \mathsf{T}_s^+(A)$, we get

$$\begin{aligned}\mathsf{T}_{2s}^+(A) &\le \tfrac{4}{3}(16C_*)L^4\Delta^4\Big(\frac{M^{2s}|A|^{2s}|P|^2}{\Delta^2 p} + \frac{M^{3s/2}|A|^{3s/2}|P|^{3/2}}{\Delta^{3/2}|B|^{1/2}}\Big) + 4|A|^{2s-1}\mathsf{T}_s^+(A)\\
&= \tfrac{4}{3}(16C_*)L^4\Big(\frac{M^{2s}|A|^{2s}|P|^2\Delta^2}{p} + \frac{M^{3s/2}|A|^{3s/2}|P|^{3/2}\Delta^{5/2}}{|B|^{1/2}}\Big) + 4|A|^{2s-1}\mathsf{T}_s^+(A)\\
&\le \tfrac{4}{3}(16C_*)L^4\Big(\frac{M^{2s}|A|^{4s}}{p} + \frac{M^{3s/2}|A|^{3s/2}(|P|\Delta^2)(|P|\Delta)^{1/2}}{|B|^{1/2}}\Big) + 4|A|^{2s-1}\mathsf{T}_s^+(A)\\
&\le 32C_*L^4\Big(\frac{M^{2s}|A|^{4s}}{p} + \frac{M^{3s/2}|A|^{2s}\mathsf{T}_s^+(A)}{|B|^{1/2}}\Big),\end{aligned}$$

and inequality (62) is proved. Here, we have used a trivial inequality $|B|^{1/2} \le |A|$ which follows from $|B| \le |AB| \le M|A| \le |B|^{1/2}|A|$ because $M^2 \le 2^{16k}M^{2^{k+1}}C_*^2\log^8|A| \le |B|$.

Now applying formula (62) successively $k-1$ times, we obtain

$$\begin{aligned}&\mathsf{T}_{2^k}(A)\\
&\le 2^{4k+6}C_*\log^4|A|\cdot\frac{M^{2^k}|A|^{2^{k+1}}}{p} + 16^{k^2}M^{2^{k+1}}C_*^{k-1}\log^{4(k-1)}|A|\cdot|A|^{2^{k+1}-4}|B|^{-\frac{k-1}{2}}\mathsf{E}^+(A), \quad (65)\end{aligned}$$

where the exponent $2^{k+1}-4$ comes from the sum $2^k+\dots+4$; to get the first term on the right-hand side of (65), we used $2^{16k}M^{2^{k+1}}C_*^2 \le |B|$ to ensure that $|B|^{1/2} \ge 2^{4k+1}C_*M^{2^k}\log^4|A|$. □

Remark 24. It is easy to see that instead of the assumption $|AB| \ll |A|$ we can assume a weaker condition $|A^s\cdot\Delta_s(B)| \ll |A|^s$, $1 < s \le 2^{k-1}$; see (64).

The same arguments work in the case of real numbers. In this situation we have no characteristic p and hence we have no any restrictions on the parameter k.

Theorem 25. *Let $A, B \subset \mathbb{R}$ be finite sets, $M \ge 1$ be a real number, and $|AB| \le M|A|$. Then for any $k \ge 2$, one has*

$$\mathsf{T}_{2^k}^+(A) \le 16^{k^2}C_*^{k-1}M^{\frac{3}{2}(2^k-1)}\log^{4(k-1)}|A|\cdot|A|^{2^{k+1}-1}|B|^{-k/2}. \quad (66)$$

Corollary 26. *Let $A \subset \mathbb{R}$ be a finite set, $M \ge 1$ be a real number, and $|AA| \le M|A|$ or $|A/A| \le M|A|$. Then for any $k \ge 2$, one has*

$$|2^kA| \gg_k |A|^{1+k/2}M^{-3/2(2^k-1)}\cdot\log^{-4(k-1)}|A|. \quad (67)$$

Bounds of such a sort were obtained in [Konyagin 2014] by another method. The best results concerning lower bounds for multiple sum sets kA, $k \to \infty$ of sets A with a small product/quotient set can be found in [Bush and Croot 2014].

To obtain an analogue of Theorem 18 for sets with $|AA| \ll |A|$, we cannot use the same arguments as in Section 4 because the spectrum is not an invariant set in this case. Moreover, in $\mathbb{R}$ there is an additional difficulty with using Fourier transform: the dual group of $\mathbb{R}$ does not coincide with $\mathbb{R}$ of course. That is why we suggest another method which works in "physical space" but not in the dual group.

To formulate our main result about $\mathsf{E}^+_k(Q)$ for sets Q with small product $Q\Gamma$ for some relatively large set Γ we need some notation. Let us write $Q^{(k)} = |Q\Gamma^{k-1}|$ for $k \ge 1$ and $Q^{(k)} = |Q|$ for $k = -1$.

Theorem 27. *Let $\Gamma, Q \subseteq \mathbb{F}_p$ be two sets, and $k \ge 0$ be an integer. Suppose that $|Q\Gamma^{k+1}||Q\Gamma^k||\Gamma| \le p^2$; further $Q^{(k)}|\Gamma| \le p$, and $M = |Q\Gamma^{k+1}|/|Q|$. Then either*

$$\mathsf{E}^+_{2^{k+1}}(Q) \le (M^{2^k+1}2^{3k+1}C_*^{(k+4)/4}\log^k Q^{(k)})\cdot|Q|^{2^{k+1}+1}|\Gamma|^{-k/8-1/2} \qquad (68)$$

or

$$\mathsf{E}^+_{2^{k+1}}(Q) \le 2(Q^{(k)})^{2^{k+1}}.$$

In particular, if we choose k such that $|\Gamma|^{k/8+1/2} \ge |Q|\cdot M^{2^k+1}2^{3k+1}C_*^{(k+4)/4}\log^k Q^{(k)}$, then

$$\mathsf{E}^+_{2^{k+1}}(Q) \le 2(Q^{(k)})^{2^{k+1}}. \qquad (69)$$

Proof. Without loss of generality one can assume that $0 \notin \Gamma$. Fix an integer $l \ge 1$ and prove that either

$$\mathsf{E}^+_{5l/2}(Q) \le 8C_*^{1/4}\log|Q|\cdot|Q|^{l/2}\mathsf{E}^+_{2l}(Q\Gamma)|\Gamma|^{-1/8} \qquad (70)$$

or

$$\mathsf{E}^+_{5l/2}(Q) \le 2|Q|^{5l/2}. \qquad (71)$$

Put $g(x) = r^l_{Q-Q}(x)$, $L = \log|Q|$, and $\mathsf{E}'_{5l/2}(Q) = \mathsf{E}^+_{5l/2}(Q) - |Q|^{5l/2} \ge 0$. We will assume below that $\mathsf{E}'_{5l/2}(Q) \ge 2^{-1}\mathsf{E}^+_{5l/2}(Q)$ because otherwise we obtain (71) immediately. Using the Dirichlet principle, we find a set P and a positive number Δ such that $P = \{x : \Delta < g(x) \le 2\Delta\} \subseteq \mathbb{F}_p^*$ and

$$\mathsf{E}'_{5l/2}(Q) \le L\sum_{x\in P} r^{5l/2}_{Q-Q}(x).$$

Applying Corollary 11, we obtain

$$\mathsf{E}'_{5l/2}(Q) \le L(2\Delta)^{3/2}\sum_{x\in P} r^l_{Q-Q}(x) \le 3C_*^{1/4}L\Delta^{3/2}|Q|^{l/2}(\mathsf{E}^+_{2l}(Q\Gamma))^{1/4}\left(\frac{|P|^4}{p}+\frac{|P|^3}{|\Gamma|^{1/2}}\right)^{1/4}$$
$$\le 3C_*^{1/4}L|Q|^{l/2}(\mathsf{E}^+_{2l}(Q\Gamma))^{1/4}\left(\frac{\Delta^6|P|^4}{p}+\frac{\Delta^6|P|^3}{|\Gamma|^{1/2}}\right)^{1/4}.$$

We have $\Delta|P| \le \mathsf{E}^+_l(Q)$, $\Delta^2|P| \le \mathsf{E}^+_{2l}(Q)$ and hence $\Delta^6|P|^4 \le (\mathsf{E}^+_{2l}(Q))^2(\mathsf{E}^+_l(Q))^2$. It follows that

$$\mathsf{E}'_{5l/2}(Q) \le 3C_*^{1/4}L|Q|^{l/2}(\mathsf{E}^+_{2l}(Q\Gamma))^{1/4}\left(\frac{(\mathsf{E}^+_{2l}(Q))^2(\mathsf{E}^+_l(Q))^2}{p}+\frac{(\mathsf{E}^+_{2l}(Q))^3}{|\Gamma|^{1/2}}\right)^{1/4}.$$

To prove that the first term $(\mathsf{E}^+_{2l}(Q))^2(\mathsf{E}^+_l(Q))^2/p$ is less than $(\mathsf{E}^+_{2l}(Q))^3/|\Gamma|^{1/2}$, we need to check that

$$(\mathsf{E}^+_l(Q))^2|\Gamma|^{1/2} \le \mathsf{E}^+_{2l}(Q)p.$$

But using the Hölder inequality, we see that the required estimate follows from

$$(\mathsf{E}^+_l(Q))^2|\Gamma|^{1/2} \le (\mathsf{E}^+_{2l}(Q))^{\frac{2(l-1)}{2l-1}}|Q|^{\frac{4l}{2l-1}}|\Gamma|^{1/2} \le \mathsf{E}^+_{2l}(Q)p$$

or, in other words, from

$$|Q|^{4l}|\Gamma|^{(2l-1)/2} \le \mathsf{E}^+_{2l}(Q)p^{2l-1}. \qquad (72)$$

Finally, we can suppose that for any $s \ge 2$ one has, say,

$$\mathsf{E}^+_s(Q) \ge |Q|^{s+1}|\Gamma|^{-1/8\log s - 1/2},$$

because otherwise estimate (68) follows easily. Our assumption $Q^{(k)}|\Gamma| \le p$ implies that $|Q||\Gamma| \le p$ and whence

$$|Q|^{2l-1}|\Gamma|^{\frac{1}{8}\log 2l + l} \le p^{2l-1}|\Gamma|^{1+1/8\log 2l - l} \le p^{2l-1},$$

and thus (72) takes place for $l \ge 2$. For $l = 1$, see the calculations below. Hence under this assumption and the inequality $\mathsf{E}'_{5l/2}(Q) \ge 2^{-1}\mathsf{E}^+_{5l/2}(Q)$, we have

$$\mathsf{E}^+_{5l/2}(Q) \le 8C_*^{1/4}\log|Q|\cdot|Q|^{l/2}\mathsf{E}^+_{2l}(Q\Gamma)|\Gamma|^{-1/8}$$

and we have proved (70). Trivially, it implies that

$$\mathsf{E}^+_{4l}(Q) \le 8C_*^{1/4}\log|Q|\cdot|Q|^{2l}\mathsf{E}^+_{2l}(Q\Gamma)|\Gamma|^{-1/8}$$

and subsequently using this bound, we obtain

$$\begin{aligned}\mathsf{E}^+_{2^k+1}(Q) &\le (2^{3k}C_*^{k/4}\log^k|Q\Gamma^{k-1}|)\cdot M^{2^{k-1}+\cdots+2}|Q|^{2^k+\cdots+2}\mathsf{E}^+(Q\Gamma^k)|\Gamma|^{-k/8}\\ &= (2^{3k}M^{2^k-2}C_*^{k/4}\log^k|Q\Gamma^{k-1}|)\cdot|Q|^{2^{k+1}-2}\mathsf{E}^+(Q\Gamma^k)|\Gamma|^{-k/8}.\end{aligned}$$

At the last step, we need to check $|Q\Gamma^{k-1}||\Gamma| \le p$, and it is guaranteed by our assumption $Q^{(k)}|\Gamma| \le p$ (for $k = -1$ we just need $|Q||\Gamma| \le p$). Now recalling the assumption $|Q\Gamma^{k+1}||Q\Gamma^k||\Gamma| \le p^2$ and applying Lemma 9, we get

$$\mathsf{E}^+_{2^k+1}(Q) \le (M^{2^k+1}2^{3k+1}C_*^{(k+4)/4}\log^k|Q\Gamma^{k-1}|)\cdot|Q|^{2^{k+1}+1}|\Gamma|^{-k/8-1/2}.$$

In particular, this final step covers the remaining case $l = 1$ above. □

Remark 28. Let Γ be a multiplicative subgroup and $Q\Gamma = Q$. Then by Theorem 27 if $|Q||\Gamma| \le p$ and a number k_1 is chosen as $|\Gamma|^{k_1/8+1/2} \ge |Q|\log^{k_1}|Q|$, then $\mathsf{E}^+_{2^{k_1}+1}(Q) \ll_{k_1} |Q|^{2^{k_1}+1}$. Let us compare this with Theorem 18. By the second part of this result (see condition (38)), choosing k_2 such that $|\Gamma|^{(k_2+2)/2} \ge |Q|\log^{4k_2}|Q|$, we get $\mathsf{E}^+_{2^{k_2}+1}(Q) \ll_{k_2} |Q|^{2^{k_2}+1}|\Gamma|^{1/2}$. After that applying the second part of Corollary 20 $n := 2^{k_2+1}$ times, we obtain

$$\begin{aligned}\mathsf{E}^+_{2^{2k_2+2}}(Q) &\ll_{k_2} |Q|^{2^{2k_2+2}} + \mathsf{E}^+_{2^{k_2}+1}(Q)(|Q||\Gamma|^{-1/4\cdot 2^{-k_2}})^n\\ &\ll_{k_2} |Q|^{2^{2k_2+2}} + |Q|^{2^{k_2}+1}|\Gamma|^{1/2}|Q|^n|\Gamma|^{-1/2} \ll |Q|^{2^{2k_2+2}}.\end{aligned}$$

Thus, Theorem 18 gives a slightly better bound (in the case of multiplicative subgroups), but of the same form.

Remark 29. From formula (40), it follows that for any l one has $\mathsf{E}^+_l(Q) \geq |Q|^{2l}/p^{l-1}$. Hence the upper bound (69) has a place just for small sets Q. For example, taking the smallest possible $l = 2$ and comparing $|Q|^2$ with $|Q|^4/p$ we see that the condition $|Q| < \sqrt{p}$ is enough. If $Q = Q\Gamma$, where Γ is a multiplicative subgroup, then it is possible to refine this condition because in the proof of Theorem 18 another method (the Fourier approach) was used. We did not make such calculations.

Now we can obtain analogues of Corollaries 20 and 22.

Corollary 30. *Let $\Gamma, Q_1, Q_2 \subseteq \mathbb{F}_p$ be sets. Take $k \geq 0$ such that for $j = 1, 2$, one has*

$$|Q_j\Gamma^{k+2}||Q_j\Gamma^{k+1}||\Gamma| \leq p^2, \quad |Q_j\Gamma^k||\Gamma| \leq p, \quad |Q_j\Gamma| \leq M_*|Q_j|, \quad |Q_j\Gamma^{k+2}| \leq M|Q_j|,$$

and

$$|\Gamma|^{k/8+1/2} \geq |Q_j| \cdot M_* M^{2^k+1} 2^{3k+1} C_*^{(k+4)/4} \log^k |Q_j\Gamma^k|. \tag{73}$$

Then, for any $x \neq 0$,

$$|Q_1 \cap (Q_2 + x)| \leq 2M_* M \sqrt{|Q_1||Q_2|} \cdot |\Gamma|^{-1/2\,(2^{-k})}. \tag{74}$$

Proof. Denote by ρ the quantity $|Q_1 \cap (Q_2 + x)|$. On the one hand, applying the Cauchy–Schwarz inequality and the second part of Theorem 27 for sets ΓQ_1 and ΓQ_2, we obtain

$$\begin{aligned}\sum_y r^{2^k+1}_{\Gamma Q_1 - \Gamma Q_2}(y) &\leq (\mathsf{E}^+_{2^k+1}(\Gamma Q_1))^{1/2} (\mathsf{E}^+_{2^k+1}(\Gamma Q_2))^{1/2} \\ &\leq 2^{3k+2} M^{2^k+1} (|Q_1\Gamma||Q_2\Gamma|)^{2^k} \leq 2^{3k+2} M^{2^k+1} M_*^{2^k+1} (|Q_1||Q_2|)^{2^k}.\end{aligned}$$

On the other hand, it is easy to see that for any $y \in \Gamma x$ one has $r_{\Gamma Q_1 - \Gamma Q_2}(y) \geq \rho$. Thus,

$$\rho^{2^k+1} |\Gamma| \leq 2^{3k+2} M^{2^k+1} M_*^{2^k+1} (|Q_1||Q_2|)^{2^k},$$

and hence

$$\rho \leq 2M_* M \sqrt{|Q_1||Q_2|} \cdot |\Gamma|^{-1/2(2^{-k})}.$$

Here we have used the inequality $k \geq 5$, which easily follows from $|\Gamma| \leq |Q_j\Gamma| \leq M|Q_j|$ and (73). $\square$

In the next two corollaries we show how to replace the condition $|Q\Gamma^k| \ll |Q|$ with a condition with a single multiplication, namely, $|Q\Gamma| \ll |Q|$.

Corollary 31. *Let Γ, Q be subsets of $\mathbb{F}_p$, $M \geq 1$ be a real number, $|Q\Gamma| \leq M|Q|$. Suppose that for $k \geq 1$ one has $(2M)^{k+1}|Q||\Gamma| \leq p$, and*

$$|\Gamma|^{k/8+1/2} \geq |Q| \cdot (2M)^{(k+3)2^k} C_*^{(k+4)/4} \log^k((2M)^k |Q|). \tag{75}$$

Then, for any $A \subseteq \mathbb{F}_p$,

$$|A + Q| \geq 2^{-3}|Q| \cdot \min\{|A|, 2^{-(4+k)} M^{-(k+3)} |\Gamma|^{\frac{1}{2}2^{-k}}\}, \tag{76}$$

and for any $\alpha \neq 0$,

$$|A(Q + \alpha)| \geq 2^{-3}|Q| \cdot \min\{|A|, 2^{-(4+k)} M^{-(k+3)} |\Gamma|^{\frac{1}{2}2^{-k}}\}. \tag{77}$$

Proof. Using Lemma 7, find a set $X \subseteq Q$, $|X| \geq |Q|/2$ such that, for any l,

$$|X\Gamma^l| \leq (2M)^l |X|. \qquad (78)$$

Also, notice that $|X\Gamma| \leq |Q\Gamma| \leq 2M|X|$. We apply Corollary 30 with M replacing by $(2M)^{k+2}$, $M_* = 2M$ and see that, for any $x \neq 0$,

$$|Q_1 \cap (Q_2 + x)| \leq 2^{k+4} M^{k+3} \sqrt{|Q_1||Q_2|} \cdot |\Gamma|^{-1/2\,(2^{-k})}.$$

Here, $Q_1 = X$ and $Q_2 = X$ or $Q_2 = \alpha X$. We will check the condition $|Q_j\Gamma^{k+2}||Q_j\Gamma^{k+1}||\Gamma| \leq p^2$ of Corollary 30 later and notice that the assumptions $|Q_j\Gamma^k||\Gamma| \leq p$, $|Q_j\Gamma| \leq M_*|Q_j|$, $|Q_j\Gamma^{k+2}| \leq M|Q_j|$ easily follow from (78) and our condition $(2M)^{k+1}|Q||\Gamma| \leq p$. Now using the arguments from Corollary 22, we estimate the energies $\mathsf{E}^+(A, X)$, $\mathsf{E}^\times(A, X+\alpha)$. In particular, we obtain lower bounds for the sum set from (76) and the product set from (77). It remains to check condition $(2M)^{2k+3}|Q|^2|\Gamma| \leq p^2$. But it follows from $(2M)^{k+1}|Q||\Gamma| \leq p$ if $M \leq |\Gamma|/2$. The last inequality is a simple consequence of (75). $\square$

Now we prove an analogue of Corollary 30 where we require that $|Q_j\Gamma|$, $j = 1, 2$ are small comparable to $|Q_j|$. For simplicity, we formulate the next corollary in the situation $|Q'| = |Q|$, but of course the general bound takes place as well.

Corollary 32. *Let Γ, Q, Q' be subsets of $\mathbb{F}_p$, $|Q'| = |Q|$, $M \geq 1$ be a real number, $|Q\Gamma|, |Q'\Gamma| \leq M|Q|$. Suppose that for $k \geq 1$ one has $(2M)^{k+1}|Q||\Gamma| \leq p$, and*

$$|\Gamma|^{\frac{k}{8} + \frac{1}{2(k+4)}} \geq |Q| \cdot M^{(k+3)2^k} C_*^{(k+4)/4} \log^k(|\Gamma|^{\frac{k}{2(k+4)} 2^{-k}} |Q|)$$

Then for any $x \neq 0$ one has

$$|Q \cap (Q' + x)| \leq 4M|Q| \cdot |\Gamma|^{-\frac{1}{2(k+4)} 2^{-k}}. \qquad (79)$$

Proof. Let $\tilde{Q} = Q \cap (Q' + x)$. Then $|\tilde{Q}\Gamma| \leq |Q\Gamma| \leq M|Q| = M|Q|/|\tilde{Q}| \cdot |\tilde{Q}| := \tilde{M}|\tilde{Q}|$. Similarly, $|(\tilde{Q} - x)\Gamma| \leq |Q'\Gamma| \leq M|Q|$. Applying the second part of Corollary 31 with $\alpha = x$, $Q = \tilde{Q}$, $A = \Gamma$, and $M = \tilde{M}$, we get

$$M|Q| \geq |(\tilde{Q} - x)\Gamma| \geq 2^{-(7+k)} |\tilde{Q}| \tilde{M}^{-(k+3)} |\Gamma|^{\frac{1}{2} 2^{-k}} = 2^{-(7+k)} M^{-(k+3)} |\tilde{Q}|^{k+4} |Q|^{-(k+3)} |\Gamma|^{\frac{1}{2} 2^{-k}}$$

provided

$$\begin{aligned} |\Gamma|^{k/8 + 1/2} &\geq |Q| \cdot (2\tilde{M})^{(k+3)2^k} C_*^{(k+4)/4} \log^k((2\tilde{M})^k |Q|) \\ &\geq |\tilde{Q}| \cdot (2\tilde{M})^{(k+3)2^k} C_*^{(k+4)/4} \log^k((2\tilde{M})^k |\tilde{Q}|). \end{aligned}$$

This gives us

$$|\tilde{Q}| \leq 4M|Q| \cdot |\Gamma|^{-\frac{1}{2(k+4)} 2^{-k}}. \qquad (80)$$

Now if the last inequality does not hold, then

$$M|Q|/\tilde{M} = |\tilde{Q}| \geq 4M|Q| \cdot |\Gamma|^{-\frac{1}{2(k+4)} 2^{-k}}$$

and thus $\tilde{M} \leq |\Gamma|^{\frac{1}{2(k+4)}2^{-k}}/4$. Hence the condition

$$|\Gamma|^{\frac{k}{8}+\frac{1}{2(k+4)}} \geq |Q| \cdot M^{(k+3)2^k} C_*^{(k+4)/4} \log^k (|\Gamma|^{\frac{k}{2(k+4)}2^{-k}}|Q|)$$

is enough. □

Now we are ready to prove the main asymmetric sum-product result of this section.

Corollary 33. *Let $A, B, C \subseteq \mathbb{F}_p$ be arbitrary sets, and $k \geq 1$ be such that $|A||B|^{1+\frac{k+1}{2(k+4)}2^{-k}} \leq p$ and*

$$|B|^{\frac{k}{8}+\frac{1}{2(k+4)}} \geq |A| \cdot C_*^{(k+4)/4} \log^k (|A||B|). \tag{81}$$

Then

$$\max\{|AB|, |A+C|\} \geq 2^{-3}|A| \cdot \min\{|C|, |B|^{\frac{1}{2(k+4)}2^{-k}}\}, \tag{82}$$

and for any $\alpha \neq 0$,

$$\max\{|AB|, |(A+\alpha)C|\} \geq 2^{-3}|A| \cdot \min\{|C|, |B|^{\frac{1}{2(k+4)}2^{-k}}\}. \tag{83}$$

Moreover,

$$|AB| + \frac{|A|^2|C|^2}{\mathsf{E}^+(A,C)} \geq 2^{-4}|A| \cdot \min\{|C|, |B|^{\frac{1}{4(k+4)}2^{-k}}\}, \tag{84}$$

and, for any $\alpha \neq 0$, we have

$$|AB| + \frac{|A|^2|C|^2}{\mathsf{E}^\times(A+\alpha,C)} \geq 2^{-4}|A| \cdot \min\{|C|, |B|^{\frac{1}{4(k+4)}2^{-k}}\}, \tag{85}$$

provided

$$|B|^{\frac{k}{8}-1/4+\frac{1}{4(k+4)}} \geq |A| \cdot C_*^{(k+4)/4} \log^k (|A||B|).$$

Proof. We will prove just (82) because the same arguments hold for (83). Put $|AB| = M|A|$, $M \geq 1$, and apply Corollary 31 with $Q = A$, $\Gamma = B$, $A = C$. Supposing that

$$|B|^{k/8+1/2} \geq |A| \cdot 2^{(k+3)2^k} M^{(k+3)2^k} C_*^{(k+4)/4} \log^k ((2M)^k|A|), \tag{86}$$

we obtain

$$|A+C| \geq 2^{-3}|A| \cdot \min\{|C|, 2^{-(k+4)}M^{-(k+3)}|B|^{\frac{1}{2}2^{-k}}\}. \tag{87}$$

Put $M_0 = 2^{-2}|B|^{\frac{1}{2(k+4)}2^{-k}}$ and consider two cases: $M \geq M_0$ and $M < M_0$. If $M \geq M_0$, then there is nothing to prove. If not, then we apply (87) and obtain the same. In other words,

$$\max\{|AB|, |A+C|\} \geq 2^{-3}|A| \cdot \min\{|C|, |B|^{\frac{1}{2(k+4)}2^{-k}}\}.$$

To check (86), we use $M < M_0$ and see that the inequality

$$|B|^{k/8+1/2} \geq |A| \cdot 2^{(k+3)2^k} M_0^{(k+3)2^k} C_*^{(k+4)/4} \log^k ((2M_0)^k|A|)$$

follows from our condition (81). The condition $(2M)^{k+1}|A||B| \leq p$ gives us $|A||B|^{1+\frac{k+1}{2(k+4)}2^{-k}} \leq p$.

To prove (84), (85), we use Corollary 32 instead of Corollary 31 and apply the arguments of the proof of Corollary 22. We obtain $\mathsf{E}^+(A,C), \mathsf{E}^\times(A+\alpha,C) \leq 2|A||C| + 4|A||C|^2 \cdot M \cdot |B|^{-\frac{1}{2(k+4)}2^{-k}}$. After that it remains to compare M with the optimal value $M_0 = 2^{-1}|B|^{\frac{1}{4(k+4)}2^{-k}}$. □

Notice that one cannot obtain any nontrivial bounds for $\min\{\mathsf{E}^{\times}(A,B), \mathsf{E}^{+}(A,C)\}$. Just take B equals a geometric progression, C equals an arithmetic progression, $|B| = |C|$, and $A = B \cup C$.

Remark 34. The results of this section take place in $\mathbb{R}$. In this case we do not need in any conditions containing the characteristic p.

Corollary 33 gives us a series of examples of "superquadratic expanders" [Balog et al. 2017] with four variables, i.e., functions $f(x_1,\dots,x_n)$ such that for any finite $A \subset \mathbb{R}$ one has

$$\left|\{f(x_1,\dots,x_n) : (x_1,\dots,x_n) \in A^n\}\right| \gg |A|^{2+c},$$

where $c > 0$ is an absolute constant. The first example of such an expander with four variables was given in [Rudnev 2017a], namely the cross-ratio function

$$f(x,y,z,w) = \frac{(y-x)(w-z)}{(z-x)(w-y)}$$

(see also [Rudnev 2017b]). It would be is interesting to find an example of a rational superquadratic expander with three variables.

Corollary 35. *Let $\varphi : \mathbb{R} \to \mathbb{R}$ be an injective function. Then for any $\kappa < \frac{1}{40}2^{-16}$ and an arbitrary finite set $A \subset \mathbb{R}$, one has $|R[A]\varphi(A)| \gg |A|^{2+\kappa}$. In particular,*

$$R[A]A = \left\{\frac{(y-x)w}{z-x} : x,y,z,w \in A,\ x \neq z\right\}$$

is a superquadratic expander with four variables.

Moreover, for any finite sets A, B, C, D of equal sizes one has

$$\left|\left\{\frac{(y-x)w}{z-x} : x \in A,\ y \in B,\ z \in C,\ w \in D,\ x \neq z\right\}\right| \gg |A|^{2+\kappa}. \tag{88}$$

Proof. By a result from [Jones 2013; Roche-Newton 2015], we have $|R[A]| \gg |A|^2/\log|A|$. Further $R[A] = 1 - R[A]$ and $R^{-1}[A] = R[A]$; see Remark 21. Hence applying estimate (83) of Corollary 33 with $A = R[A]$, $B = C = \varphi(A)$, and $\alpha = -1$, we obtain

$$|R[A]\varphi(A)| \gg |R[A]| \cdot |A|^{\frac{1}{2(k+4)}2^{-k}},$$

provided

$$|A|^{\frac{k}{8}+\frac{1}{2(k+4)}} \geq |R[A]| \cdot 2^{2k} C_*^{(k+4)/4} \log^k |A| \geq |R[A]| \cdot C_*^{(k+4)/4} \log^k |R[A]\varphi(A)|. \tag{89}$$

Put $|R[A]| = C|A|^{2+c}/\log|A|$, $c \geq 0$, and $C > 0$ is an absolute constant. Then taking $k = 16 + 8c$, say, we satisfy (89) for large A. It follows that

$$|R[A]\varphi(A)| \gg |A|^{2+c+\frac{1}{2(20+8c)}2^{-16-8c}} \log^{-1}|A|.$$

One can check that the optimal choice of c is $c = 0$. Finally, to prove (88) just notice that from the method of [Jones 2013; Roche-Newton 2015] it follows that

$$\left|\left\{\frac{b-a}{c-a} : a \in A,\ b \in B,\ c \in C,\ c \neq a\right\}\right| \gg |A|^2/\log|A|$$

for any sets A, B, C of equal cardinality. After that, repeat the arguments above. □

Remark 36. Let us show quickly how Corollary 33 implies both Theorems 1, 2 for sets A with $|A| < p^{1/2-\varepsilon}$ (the appearance $\sqrt{p}$ bound was discussed in Remark 29).

Let B, C be sets of sizes greater than p^{ε} such that $\max\{|AB|, |A+C|\} \le p^{\delta}|A|$ or $\max\{|(A+\alpha)B|, |A+C|\} \le p^{\delta}|A|$ for some $\alpha \neq 0$. We can find sufficiently large $k = k(\varepsilon)$ such that condition (81) takes place for B because $|A| < p^{1/2-\varepsilon} \le p$ and $|B| \ge p^{\varepsilon}$. Applying Corollary 33 for A, B, C, we arrive to a contradiction. Finally, to ensure that $|A||B|^{1+\frac{k+1}{2(k+4)}2^{-k}} \le p$ just use the assumption $|A| < p^{1/2-\varepsilon}$, inequality $|B| \le |AB| \le p^{\delta}|A|$, and take sufficiently small $\delta = \delta(\varepsilon)$ and sufficiently large $k = k(\varepsilon)$.

Let $A \subset \mathbb{R}$ be a finite set. We consider a characteristic of A (see, e.g., [Shkredov 2016a]) that generalizes the notion of small multiplicative doubling of A. Namely, put

$$d^{+}(A) := \inf_{f} \min_{B \neq \varnothing} \frac{|f(A)+B|^2}{|A||B|},$$

where the infimum is taken over convex/concave functions f.

Problem. Suppose that $d^{+}(A) \le |A|^{\varepsilon}$ and $\varepsilon > 0$ is a small number. Is it true that there is $k = k(\varepsilon)$ such that $\mathsf{E}^{+}_{k}(A) \ll |A|^{k}$?

Notice that one cannot obtain a similar bound for $\mathsf{T}^{+}_{k}(A)$. Indeed, let $A = \{1^2, 2^2, \dots, n^2\}$. Then one can show that for such A, the quantity $d^{+}(A)$ is $O(1)$ (see, e.g., [Shkredov 2016a]) but, clearly, $|kA| \ll_k |A|^2$. This means that it is not possible to obtain any upper bound for $\mathsf{T}^{+}_{k}(A)$ of the form $\mathsf{T}^{+}_{k}(A) \ll |A|^{2k-2-c}$, $c > 0$, and hence any analogues of Theorems 23, 25 for sets A with small $d^{+}(A)$.

Acknowledgements

The author thanks Misha Rudnev and Sophie Stevens for careful reading of the first draft of this paper and for useful discussions. Also he thanks the referee for valuable suggestions, remarks and careful reading of our article. This work is supported by the Russian Science Foundation under grant 14-11-00433.

References

[Balog et al. 2017] A. Balog, O. Roche-Newton, and D. Zhelezov, "Expanders with superquadratic growth", *Electron. J. Combin.* **24**:3 (2017), art. id. 3.14. MR Zbl

[Bourgain 2003] J. Bourgain, "On the Erdős–Volkmann and Katz–Tao ring conjectures", *Geom. Funct. Anal.* **13**:2 (2003), 334–365. MR Zbl

[Bourgain 2005a] J. Bourgain, "Estimates on exponential sums related to the Diffie–Hellman distributions", *Geom. Funct. Anal.* **15**:1 (2005), 1–34. MR Zbl

[Bourgain 2005b] J. Bourgain, "Exponential sum estimates over subgroups of $\mathcal{Z}_q^*$, q arbitrary", *J. Anal. Math.* **97** (2005), 317–355. MR

[Bourgain 2005c] J. Bourgain, "More on the sum-product phenomenon in prime fields and its applications", *Int. J. Number Theory* **1**:1 (2005), 1–32. MR Zbl

[Bourgain 2007] J. Bourgain, "Exponential sum estimates in finite commutative rings and applications", *J. Anal. Math.* **101** (2007), 325–355. MR Zbl

[Bourgain and Garaev 2009] J. Bourgain and M. Z. Garaev, "On a variant of sum-product estimates and explicit exponential sum bounds in prime fields", *Math. Proc. Cambridge Philos. Soc.* **146**:1 (2009), 1–21. MR Zbl

[Bourgain et al. 2004] J. Bourgain, N. Katz, and T. Tao, "A sum-product estimate in finite fields, and applications", *Geom. Funct. Anal.* **14**:1 (2004), 27–57. MR Zbl

[Bourgain et al. 2006] J. Bourgain, A. A. Glibichuk, and S. V. Konyagin, "Estimates for the number of sums and products and for exponential sums in fields of prime order", *J. London Math. Soc.* (2) **73**:2 (2006), 380–398. MR Zbl

[Bush and Croot 2014] A. Bush and E. Croot, "Few products, many h-fold sums", preprint, 2014. arXiv

[Erdős and Szemerédi 1983] P. Erdős and E. Szemerédi, "On sums and products of integers", pp. 213–218 in *Studies in pure mathematics*, edited by P. Erdős, Birkhäuser, Basel, 1983. MR

[Garaev 2010] M. Z. Garaev, "Sums and products of sets and estimates for rational trigonometric sums in fields of prime order", *Uspekhi Mat. Nauk* **65**:4 (2010), 599–658. In Russian; translated in *Russian Math. Surveys* **65**:4 (2010), 599–658. MR Zbl

[Glibichuk and Konyagin 2007] A. A. Glibichuk and S. V. Konyagin, "Additive properties of product sets in fields of prime order", pp. 279–286 in *Additive combinatorics*, edited by A. Granville et al., CRM Proc. Lecture Notes **43**, Amer. Math. Soc., Providence, RI, 2007. MR Zbl

[Jones 2013] T. G. F. Jones, "New quantitative estimates on the incidence geometry and growth of finite sets", preprint, 2013. arXiv

[Konyagin 2014] S. Konyagin, "h-fold sums from a set with few products", *Mosc. J. Comb. Number Theory* **4**:3 (2014), 14–20. MR Zbl

[Konyagin and Shkredov 2016] S. V. Konyagin and I. D. Shkredov, "New results on sums and products in $\mathcal{R}$", *Tr. Mat. Inst. Steklova* **294** (2016), 87–98. In Russian. MR Zbl

[Konyagin and Shparlinski 1999] S. V. Konyagin and I. E. Shparlinski, *Character sums with exponential functions and their applications*, Cambridge Tracts in Mathematics **136**, Cambridge University Press, 1999. MR Zbl

[Murphy et al. 2017] B. Murphy, G. Petridis, O. Roche-Newton, M. Rudnev, and I. D. Shkredov, "New results on sum-product type growth over fields", preprint, 2017. arXiv

[Roche-Newton 2015] O. Roche-Newton, "A short proof of a near-optimal cardinality estimate for the product of a sum set", pp. 74–80 in *31st International Symposium on Computational Geometry*, vol. 34, edited by L. Arge, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, Germany, 2015. MR Zbl

[Roche-Newton et al. 2016] O. Roche-Newton, M. Rudnev, and I. D. Shkredov, "New sum-product type estimates over finite fields", *Adv. Math.* **293** (2016), 589–605. MR

[Rudnev 2017a] M. Rudnev, "On distinct cross-ratios and related growth problems", preprint, 2017. arXiv

[Rudnev 2017b] M. Rudnev, "On the number of incidences between points and planes in three dimensions", *Combinatorica* (2017).

[Ruzsa 2009] I. Z. Ruzsa, "Sumsets and structure", pp. 87–210 in *Combinatorial number theory and additive group theory*, edited by M. Castellet, Birkhäuser, Basel, Switzerland, 2009. MR Zbl

[Schoen and Shkredov 2013] T. Schoen and I. D. Shkredov, "Higher moments of convolutions", *J. Number Theory* **133**:5 (2013), 1693–1737. MR Zbl

[Shkredov 2013] I. D. Shkredov, "Some new inequalities in additive combinatorics", *Mosc. J. Comb. Number Theory* **3**:3-4 (2013), 189–239. MR Zbl

[Shkredov 2014] I. Shkredov, "Energies and structure of additive sets", *Electron. J. Combin.* **21**:3 (2014), art. id. 3.44. MR Zbl

[Shkredov 2016a] I. D. Shkredov, "Difference sets are not multiplicatively closed", *Discrete Anal.* (2016), art. id. 17. MR Zbl

[Shkredov 2016b] I. D. Shkredov, "On tripling constant of multiplicative subgroups", *Integers* **16** (2016), art. id. A75. MR Zbl

[Shkredov 2017] I. Shkredov, "Some remarks on the Balog–Wooley decomposition theorem and quantities D^+, D^x", *Proc. Steklov Inst. Math.* **298** (2017), 74–90.

[Shteinikov 2015] Y. N. Shteinikov, "Estimates of trigonometric sums over subgroups and some of their applications", *Math. Notes* **98**:3-4 (2015), 606–625. MR

[Szemerédi and Trotter 1983] E. Szemerédi and W. T. Trotter, Jr., "Extremal problems in discrete geometry", *Combinatorica* **3**:3-4 (1983), 381–392. MR Zbl

[Tao and Vu 2006] T. Tao and V. Vu, *Additive combinatorics*, Cambridge Studies in Advanced Mathematics **105**, Cambridge University Press, 2006. MR Zbl

[Yazici et al. 2017] E. A. Yazici, B. Murphy, M. Rudnev, and I. Shkredov, "Growth estimates in positive characteristic via collisions", *Int. Math. Res. Not.* **2017**:23 (2017), 7148–7189.

Received 1 Dec 2017.

ILYA D. SHKREDOV:

ilya.shkredov@gmail.com
Steklov Mathematical Institute, ul. Gubkina, 9, Moscow, Russia, 119991

and

IITP RAS, Bolshoy Karetny per. 19, Moscow, Russia, 127994

and

MIPT, Institutskii per. 9, Dolgoprudnii, Russia, 141701

msp

Moscow Journal of Combinatorics and Number Theory
Vol. 8, No. 1, 2019
dx.doi.org/10.2140/moscow.2019.8.43

Convex sequences may have thin additive bases

Imre Z. Ruzsa and Dmitrii Zhelezov

For a fixed $c > 0$ we construct an arbitrarily large set B of size n such that its sum set $B + B$ contains a convex sequence of size cn^2, answering a question of Hegarty.

Notation

The following notation is used throughout the paper. The expressions $X \gg Y$, $Y \ll X$, $Y = O(X)$, $X = \Omega(Y)$ all have the same meaning in that there is an absolute constant c such that $|Y| \le c|X|$.

If X is a set then $|X|$ denotes its cardinality.

For sets of numbers A and B the *sumset* $A + B$ is the set of all pairwise sums

$$\{a + b : a \in A, b \in B\}.$$

1. Introduction

Let $A = \{a_i\}$, $i = 1, \dots, n$, be a set of real numbers ordered in a way that $a_1 \le a_2 \le \cdots \le a_n$. (We also refer to A a sequence, if we wish to emphasize the ordering.) Recall that A is called *convex* if the gaps between consecutive elements of A are strictly increasing, that is

$$a_2 - a_1 < a_3 - a_2 < \cdots < a_n - a_{n-1}.$$

Studies of convex sets were initiated by Erdős, who conjectured that any convex set must grow with respect to addition, so that the size of the set of sums $A + A := \{a_1 + a_2 : a_1, a_2 \in A\}$ is significantly larger than the size of A.

The first nontrivial bound confirming the conjecture of Erdős was obtained by Hegyvári [1986]. The state of the art bound for the size of $A + A$ of a convex sequence A is due to Shkredov [2015]:

$$|A + A| \gg |A|^{58/37} \log^{-20/37} |A|.$$

The best bound for the size of the difference set $A - A$ is due to Schoen and Shkredov [2011], who proved that

$$|A - A| \gg |A|^{8/5} \log^{-2/5} |A|$$

if A is arbitrary convex sequence. It is conjectured that in fact

$$|A + A| \ge C(\epsilon)|A|^{2-\epsilon}$$

holds for any $\epsilon > 0$ and some $C > 0$ which depends only on ϵ.

MSC2010: 11B13.

Keywords: convex sequences, sumset, additive basis.

In general, it is believed that convex sets cannot be additively structured. In particular, Hegarty [2012] asked whether there is a constant $c > 0$ with the property that there is a set B of arbitrarily large size n such that $B + B$ contains a convex set of size cn^2.

Recall that B is a *basis* (of order two) for a set A if $A \subset B + B$. In other words, Hegarty asked if a convex set of size n can have a thin additive basis (of order two) of size as small as $O(n^{1/2})$, which is clearly the smallest possible size up to a constant.

Perhaps contrary to the intuition that convex sets lack additive structure, we present a construction which answers Hegarty's question in the affirmative. Our main result is as follows.

Theorem 1. *There is $c > 0$ such that for any m there is a set B of size $n > m$ such that $B + B$ contains a convex set of size cn^2.*

2. Construction

Assume n is fixed and large. We will construct a set B of size $O(n)$ such that $B + B$ contains a convex set of size $\Omega(n^2)$. Theorem 1 will clearly follow.

The following constants (we assume n is fixed) will be used throughout the proof:

$$\alpha := \frac{1}{n^2}, \quad \gamma := \frac{1}{1000n^3}, \quad \epsilon := 0.1.$$

Define

$$x_i = i + (\alpha + \gamma)i^2, \quad y_j = j - \alpha j^2.$$

Next, we define

$$B_k = \{x_i + y_j : i + j = k\},$$

where i and j are allowed to be negative.

Let $k \in [0.999n, n]$ so that $\alpha k^2 \in [0.99, 1]$. For such an integer k writing $j = k - i$ we have that the i-th element of B_k is given by

$$b_i^{(k)} = k + (\alpha + \gamma)i^2 - \alpha(k - i)^2 = (k - \alpha k^2) + \gamma i^2 + 2ik\alpha. \tag{1}$$

Now assume that i ranges in $[-n, 2n]$. The consecutive differences $b_{i+1}^{(k)} - b_i^{(k)}$ are then given by

$$\Delta_i^{(k)} := \gamma(2i + 1) + 2k\alpha.$$

Observe that $\Delta_i^{(k)}$ are positive and increasing, thus the block $B_k := \{b_i^{(k)}\}_{-n}^{2n}$ is convex. Further, by (1) for sufficiently large n we have

$$b_{-n}^{(k)} = k - \alpha k^2 + \gamma n^2 - 2nk\alpha \in [k - 2.9, k - 3], \tag{2}$$

$$b_{2n}^{(k)} = k - \alpha k^2 + \gamma(2n)^2 + 4nk\alpha \in [k + 2.9, k + 3.1], \tag{3}$$

so $B_k \subset [k - 3, k + 3] + [-\epsilon, \epsilon]$.

Now we are going to build a large convex sequence out of blocks B_k with $4 \mid k$. Since each B_k is already convex, it remains to show how to glue together B_k and B_{k+4} so that the resulting set is again convex. We proceed with the following simple lemma.

Lemma 2. *Let $X = \{x_i\}_{i=0}^{N}$ and $Y = \{y_j\}_{j=0}^{M}$ be two convex sequences and there are indices u and v such that*

$$[x_u, x_{u+1}] \subset [y_v, y_{v+1}].$$

Then

$$Z := \{x_i\}_{i=0}^{u} \cup \{y_j\}_{j=v+1}^{M}$$

is a convex sequence.

Proof. Since $[x_u, x_{u+1}] \subset [y_v, y_{v+1}]$ we have that

$$x_u - x_{u-1} < x_{u+1} - x_u < y_{v+1} - x_u.$$

On the other hand,

$$y_{v+1} - x_u < y_{v+1} - y_v < y_{v+2} - y_{v+1}. \qquad \square$$

By Lemma 2, in order to merge B_k and B_{k+4} it suffices to find two consecutive elements $b_i^{(k)}, b_{i+1}^{(k)} \in B_k$ in between two consecutive elements $b_j^{(k+4)}, b_{j+1}^{(k+4)} \in B_{k+4}$. Define

$$\delta := \max_{i \in [-n, 2n]} \Delta_i^{(k)}, \quad \Delta := \min_{i \in [-n, 2n]} \Delta_i^{(k+4)}.$$

We have

$$\delta < 4n\gamma + 2k\alpha < \frac{2.1}{n}, \tag{4}$$

$$\Delta - \delta > 8\alpha - 10n\gamma > \frac{6}{n^2}. \tag{5}$$

Let $b_v^{(k)}$ be the least element in B_k greater than $b_{-n}^{(k+4)}$ (such an element exists by (2) and (3)). We claim that with $m := \lceil n/2 \rceil + 1$ holds $b_{-n+m}^{(k+4)} > b_{v+m}^{(k)}$, which in turn by the pigeonhole principle guarantees the arrangement of elements required by Lemma 2.

Indeed, by our choice of v,

$$0 \le d := b_v^{(k)} - b_{-n}^{(k+4)} \le \delta. \tag{6}$$

But by (4) and (5),

$$b_{-n+m}^{(k+4)} - b_{v+m}^{(k)} > -d + m(\Delta - \delta) > \frac{3}{n} - \delta > 0, \tag{7}$$

so the claim follows.

It remains to note that by (3),

$$b_{v+m}^{(k)} < b_{-n}^{(k+4)} + m\Delta < (k + 1 + \epsilon) + \frac{2n^2\alpha}{2} + 4\gamma nm < k + 2.2,$$

and thus $v + m < 2n$ again by (3). This verifies that $b_v^{(k)}, b_{v+m}^{(k)} \in B_k$.

3. Putting everything together

Applying the procedure described in the previous section, we can glue together consecutive blocks B_{4l} with $4l := k \in [0.999n, n]$. Let A be the resulting convex sequence. First, observe there are $\Omega(n)$ blocks being merged. Moreover, each interval $[4l - 1 + \epsilon, 4l + 1 - \epsilon]$ is covered only by the block B_{4l} and by (2), (3), and (4) it contains $\Omega(n)$ elements from B_{4l}, so $|A| = \Omega(n^2)$. On the other hand, by our construction, A is contained in the sumset $B + B$ of $B := \{x_i\}_{-2n}^{2n} \cup \{y_j\}_{-2n}^{2n}$ of size $O(n)$.

Remark 3. It follows from our construction that there are arbitrarily large convex sets A such that the equation

$$a_1 - a_2 = x : a_1, a_2 \in A$$

has $\Omega(|A|^{1/2})$ solutions (a_1, a_2) for at least $\Omega(|A|^{1/2})$ values of x.

Acknowledgments

Ruzsa is supported by ERC-AdG. 321104 and Hungarian National Research Development and Innovation Funds K 109789, NK 104183 and K 119528. Zhelezov is supported by the Knut and Alice Wallenberg postdoctoral fellowship.

The work on this paper was partially carried out while Zhelezov was visiting the Rényi Institute of Mathematics by invitation of Endre Szemerédi, whose hospitality and support is greatly acknowledged. We also thank Peter Hegarty and Ilya Shkredov for useful discussions.

References

[Hegarty 2012] P. Hegarty, "Convex subsets of sumsets", MathOverflow post, 2012, https://mathoverflow.net/questions/106817. The original formulation is contrapositive to ours, which is more convenient to state.

[Hegyvári 1986] N. Hegyvári, "On consecutive sums in sequences", *Acta Math. Hungar.* **48**:1-2 (1986), 193–200. MR Zbl

[Schoen and Shkredov 2011] T. Schoen and I. D. Shkredov, "On sumsets of convex sets", *Combin. Probab. Comput.* **20**:5 (2011), 793–798. MR Zbl

[Shkredov 2015] I. D. Shkredov, "On sums of Szemerédi–Trotter sets", *Tr. Mat. Inst. Steklova* **289**:1 (2015), 318–327. In Russian; translated in *Proc. Steklov Inst. Math.* **289**:1 (2015), 300–309. MR Zbl

Received 1 Dec 2017.

IMRE Z. RUZSA:

ruzsa.z.imre@renyi.mta.hu
Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, Budapest, Hungary

DMITRII ZHELEZOV:

dzhelezov@gmail.com
Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, Budapest, Hungary

msp

Moscow Journal of Combinatorics and Number Theory
Vol. 8, No. 1, 2019
dx.doi.org/10.2140/moscow.2019.8.47

Admissible endpoints of gaps in the Lagrange spectrum

Dmitry Gayfulin

For any irrational number α define the Lagrange constant $\mu(\alpha)$ by

$$\mu^{-1}(\alpha) = \liminf_{p\in\mathbb{Z},\, q\in\mathbb{N}} |q(q\alpha - p)|.$$

The set of all values taken by $\mu(\alpha)$ as α varies is called the *Lagrange spectrum* $\mathbb{L}$. An irrational α is called attainable if the inequality

$$\left|\alpha - \frac{p}{q}\right| \leqslant \frac{1}{\mu(\alpha)q^2}$$

holds for infinitely many integers p and q. We call a real number $\lambda \in \mathbb{L}$ *admissible* if there exists an irrational attainable α such that $\mu(\alpha) = \lambda$. In a previous paper we constructed an example of a nonadmissible element in the Lagrange spectrum. In the present paper we give a necessary and sufficient condition for admissibility of a Lagrange spectrum element. We also give an example of an infinite sequence of left endpoints of gaps in $\mathbb{L}$ which are not admissible.

1. Introduction

The Lagrange spectrum $\mathbb{L}$ is usually defined as the set of all values of the Lagrange constants

$$\mu(\alpha) = \Big(\liminf_{p\in\mathbb{Z},\, q\in\mathbb{N}} |q(q\alpha - p)|\Big)^{-1}$$

as α runs through the set of irrational numbers. Consider the continued fraction expansion of α

$$\alpha = [a_0; a_1, a_2, \ldots, a_n, \ldots].$$

For any positive integer i define

$$\lambda_i(\alpha) = [a_i; a_{i+1}, a_{i+2}, \ldots] + [0; a_{i-1}, a_{i-2}, \ldots, a_1].$$

It is well-known fact that

$$\limsup_{i\to\infty} \lambda_i(\alpha) = \mu(\alpha). \tag{1}$$

The equation (1) provides an equivalent definition of the Lagrange constant $\mu(\alpha)$.

The following properties of $\mathbb{L}$ are well known. The Lagrange spectrum is a closed set [Cusick 1975] with minimal point $\sqrt{5}$. All the numbers of $\mathbb{L}$ which are less than 3 form a discrete set. The Lagrange spectrum contains all elements greater than $\sqrt{21}$; see [Freiman 1973; Schecker 1977]. The complement

Research supported by RNF grant No. 14-11-00433. The author is a Young Russian Mathematics award winner and would like to thank its sponsors and jury.

MSC2010: 11J06.

Keywords: Lagrange spectrum, Diophantine approximation, continued fractions.

of $\mathbb{L}$ is a countable union of *maximal gaps* of the spectrum. The maximal gaps are open intervals (a, b) such that $(a, b) \cap \mathbb{L} = \varnothing$, but a and b both lie in the Lagrange spectrum. There are infinitely many gaps in the nondiscrete part of the Lagrange spectrum [Gbur 1976].

Let α be an arbitrary irrational number. If the inequality

$$\left|\alpha - \frac{p}{q}\right| \leqslant \frac{1}{\mu(\alpha)q^2}$$

has infinitely many solutions for integer p and q, we call α *attainable*. This definition was first given in [Malyshev 1977]. One can easily see [Gayfulin 2017] that α is attainable if and only if $\lambda_i(\alpha) \geqslant \mu(\alpha)$ for infinitely many indices i. We also call a real number $\lambda \in \mathbb{L}$ *admissible* if there exists an irrational attainable number α such that $\mu(\alpha) = \lambda$.

Let B denote a doubly infinite sequence of positive integers

$$B = (\ldots, b_{-n}, \ldots, b_{-1}, b_0, b_1, \ldots, b_n, \ldots).$$

For an arbitrary integer i define

$$\lambda_i(B) = [b_i; b_{i-1}, \ldots] + [0; b_{i+1}, b_{i+2}, \ldots].$$

We will call a doubly infinite sequence B purely periodic if there exists a finite sequence P such that $B = (\overline{P})$. A doubly infinite sequence B is called eventually periodic if there exist three finite sequences P_l, R, P_r such that $B = (\overline{P}_l, R, \overline{P}_r)$. One can also consider an equivalent definition of the Lagrange spectrum using the doubly infinite sequences. We use the notation from [Cusick and Flahive 1989]:

$$L(B) = \limsup_{i\to\infty} \lambda_i(B), \quad M(B) = \sup \lambda_i(B).$$

The Lagrange spectrum $\mathbb{L}$ is exactly the set of values taken by $L(B)$ as B runs through the set of doubly infinite sequences of positive integers. The set of values taken by $M(B)$ is called the Markoff spectrum. We will denote this set by $\mathbb{M}$.

We will call a doubly infinite sequence B *weakly associated* with an irrational number $\alpha = [a_0; a_1, \ldots, a_n, \ldots]$ if the following condition holds:

(1) For any natural i the pattern $(b_{-i}, b_{-i+1}, \ldots, b_0, \ldots, b_i)$ occurs in the sequence $a_1, a_2, \ldots, a_n, \ldots$ infinitely many times.

We will call B *strongly associated* with α if, additionally,

(2) $\mu(\alpha) = \lambda_0(B) = M(B)$.

One can easily see that if B is weakly associated with α then $\mu(\alpha) \geqslant M(B)$. As we will show in Lemma 4.1, if α has bounded partial quotients, it has at least one strongly associated sequence.

2. Results of [Gayfulin 2017]

Theorem I. *The quadratic irrationality* $\lambda_0 = [3; 3, 3, 2, 1, \overline{1, 2}] + [0; 2, 1, \overline{1, 2}]$ *belongs to* $\mathbb{L}$, *but if* α *is such that* $\mu(\alpha) = \lambda_0$ *then* α *is not attainable.*

Theorem II. *If* $\lambda \in \mathbb{L}$ *is not a left endpoint of some maximal gap in the Lagrange spectrum then there exists an attainable* α *such that* $\mu(\alpha) = \lambda$.

One can easily formulate these theorems using the concept of admissible numbers, introduced above.

Theorem I′. *The quadratic irrationality* $\lambda_0 = [3; 3, 3, 2, 1, \overline{1, 2}] + [0; 2, 1, \overline{1, 2}]$ *belongs to* $\mathbb{L}$, *but is not admissible.*

Theorem II′. *If* $\lambda \in \mathbb{L}$ *is not a left endpoint of some maximal gap in the Lagrange spectrum then* λ *is an admissible number.*

3. Main results

Our first theorem is a small generalization of Theorem 3 in [Gayfulin 2017]. The proof will be quite similar and use some lemmas from that paper.

Theorem 1. *Let* a *be a left endpoint of a gap* (a, b) *in the Lagrange spectrum and* α *be an irrational number such that* $\mu(\alpha) = a$. *Consider a doubly infinite sequence* B *strongly associated with* α. *Then* B *is an eventually periodic sequence.*

It follows from Theorems I and II that there exist nonadmissible elements in the Lagrange spectrum but all such numbers are left endpoints of some maximal gaps in $\mathbb{L}$. The following theorem gives a necessary and sufficient condition of admissibility of a Lagrange spectrum element.

Theorem 2. *A left endpoint of a gap in the Lagrange spectrum* a *is admissible if and only if there exists a quadratic irrationality* α *such that* $\mu(\alpha) = a$.

Of course, every quadratic irrationality is strongly associated with the unique sequence, which is purely periodic. Therefore Theorem 2 is equivalent to the following statement.

Corollary 3.1. *A left endpoint of a gap in the Lagrange spectrum* a *is not admissible if and only if there does not exist a purely periodic sequence* B *such that* $\lambda_0(B) = M(B) = a$.

Theorem 2 provides an instrument to verify nonadmissible points in $\mathbb{L}$. Define

$$\alpha_n^* = 2 + [0; \underbrace{1, \dots, 1}_{2n-2}, \overline{2, 2, 1, 2}] + [0; \underbrace{1, \dots, 1}_{2n-1}, 2, \underbrace{1, \dots, 1}_{2n-2}, \overline{2, 2, 1, 2}],$$
$$\beta_n = 2 + 2[0; \underbrace{\overline{1, \dots, 1}}_{2n}, 2].$$

The fact that (α_n^*, β_n) is the maximal gap in the Markoff spectrum was proved in [Gbur 1976]. It is easy to show that α_n^* and β_n belong to $\mathbb{L}$; we will do this in Section 6. Hence, as $\mathbb{L} \subset \mathbb{M}$ [Cusick 1975], the interval (α_n^*, β_n) is the maximal gap in $\mathbb{L}$ too.

Theorem 3. *For any integer* $n \geqslant 2$ *the irrational number* α_n^* *is not admissible.*

One can easily see that $\alpha_1^* = 2 + [0; \overline{2, 2, 1, 2}] + [0; 1, 2, \overline{2, 2, 1, 2}] = \mu([0; \overline{2, 2, 1, 2}]) = M(\overline{2, 2, 1, 2})$. Thus, α_1^* is an admissible number by Theorem 2.

4. Proof of Theorem 1

The following statement is well known. See the proof in [Cusick and Flahive 1989, Chapter 1, Lemma 6].

Lemma 4.1. *Let $A = \ldots, a_{-1}, a_0, a_1, \ldots$ be any doubly infinite sequence. If $M(A)$ is finite, then there exists a doubly infinite sequence B such that $M(A) = M(B) = \lambda_0(B)$.*

Using the same argument for the sequence $A = (a_1, a_2, \ldots, a_n, \ldots)$, one can easily show:

Lemma 4.2. *Let $\alpha = [0; a_1, \ldots, a_n, \ldots]$ be an arbitrary irrational number and $a_i < c$ for all $i \in \mathbb{N}$, for some positive real number c. Then there exists a doubly infinite sequence B which is strongly associated with α.*

As $\alpha \leqslant \sqrt{21}$, all elements of B are bounded by 4. For any natural n define $\varepsilon_n = 2^{-(n-1)}$, $\delta_n = 5^{-2(n+2)}$. We need the following lemmas from [Gayfulin 2017].

Lemma 4.3. *Suppose $\alpha = [a_0; a_1, \ldots, a_n, b_1, \ldots]$ and $\beta = [a_0; a_1, \ldots, a_n, c_1, \ldots]$, where $n \geqslant 0$, a_0 is an integer, and $a_1, \ldots, a_n, b_1, b_2, \ldots, c_1, c_2, \ldots$ are positive integers bounded by 4 with $b_1 \neq c_1$. Then for n odd, $\alpha > \beta$ if and only if $b_1 > c_1$; for n even, $\alpha > \beta$ if and only if $b_1 < c_1$. Also,*

$$\delta_n < |\alpha - \beta| < \varepsilon_n.$$

Lemma 4.4. *Let $\gamma = [0; c_1, c_2, \ldots, c_N, \ldots]$ and $\gamma' = [0; c'_1, c'_2, \ldots, c'_N, \ldots]$ be two irrational numbers with partial quotients not exceeding 4. Suppose that every sequence of partial quotients of length $2n+1$ which occurs in the sequence $(c'_1, c'_2, \ldots, c'_N, \ldots)$ infinitely many times also occurs in the sequence $(c_1, c_2, \ldots, c_N, \ldots)$ infinitely many times. Then $\mu(\gamma') < \mu(\gamma) + 2\varepsilon_n$.*

The following technical lemma was formulated in [Gayfulin 2017] for $N = (2n+1)(4^{2n+1}+1)$ and the proof was incorrect. However, this is not crucial for the results of that paper, as we just need N to be bounded from above by some growing function of n. In this paper, we give a new version of the lemma with correct proof.

Lemma 4.5. *Let n be an arbitrary positive integer. Define $N = N(n) = (2n+2)(4^{2n+2}+1)$. If $b_1, b_2, \ldots, b_N$ is an arbitrary integer sequence of length N such that $1 \leqslant b_i \leqslant 4$ for all $1 \leqslant i \leqslant N$, then there exist two integers n_1, n_2 such that $b_{n_1+i} = b_{n_2+i}$ for all $0 \leqslant i \leqslant 2n+1$ and $n_1 \equiv n_2 \pmod 2$.*

Proof. There exist only 4^{2n+2} distinct sequences of length $2n+2$ with elements $1, 2, 3, 4$. Consider $4^{2n+2}+1$ sequences: $(b_1, \ldots, b_{2n+2})$, $(b_{2n+3}, \ldots, b_{4n+4})$, $\ldots$, $(b_{(2n+2)4^{2n+2}+1}, \ldots, b_{(2n+2)4^{2n+2}+2n+2})$. Dirichlet's principle implies that there exist two coinciding sequences among them. Denote these sequences by $(b_{n_1}, \ldots, b_{n_1+2n+1})$ and $(b_{n_2}, \ldots, b_{n_2+2n+1})$. Note that the index of the first element of each sequence is odd; hence $n_1 \equiv n_2 \equiv 1 \pmod 2$, which finishes the proof. □

If $n_1 \equiv n_2 \pmod 2$ then the sequence $(b_{n_1}, b_{n_1+1}, \ldots, b_{n_2-1})$ has even length. This fact will be useful in our argument.

Lemma 4.6. *Let B be an arbitrary integer sequence of even length. Let A be an arbitrary finite integer sequence and C an arbitrary nonperiodic infinite sequence. Then*

$$\min([0; A, B, B, C], [0; A, C]) < [0; A, B, C] < \max([0; A, B, B, C], [0; A, C]). \tag{2}$$

Proof. As the sequence C is nonperiodic, the continued fractions in (2) are not equal. Without loss of generality, one can say that the sequence A is empty. Suppose that

$$[0; B, C] > [0; B, B, C].$$

As the length of B is even, one can see that $[0; C] > [0; B, C]$, which is exactly the right-hand side of (2). The case when $[0; B, C] < [0; B, B, C]$ is treated in exactly the same way. □

Lemma 4.7. *Let $\gamma = [0; b_1, b_2, \dots, b_N, \dots]$ be an arbitrary irrational number, not a quadratic irrationality. Consider the sequence $B_N = (b_1, b_2, \dots, b_N)$ and define two numbers n_1 and n_2 from Lemma 4.5. Define two new sequences of positive integers*

$$B_N^1 = (b_1, b_2, \dots, b_{n_1-1}, b_{n_2}, b_{n_2+1}, \dots, b_N),$$
$$B_N^2 = (b_1, b_2, \dots, b_{n_1-1}, b_{n_1}, \dots, b_{n_2-1}, b_{n_1}, \dots, b_{n_2-1}, b_{n_2}, b_{n_2+1}, \dots, b_N).$$

Let us also define two new irrational numbers:

$$\gamma^1 = [0; b_1, b_2, \dots, b_{n_1-1}, b_{n_2}, b_{n_2+1}, \dots, b_N, b_{N+1}, \dots] = [0; B_N^1, b_{N+1}, \dots],$$
$$\gamma^2 = [0; b_1, b_2, \dots, b_{n_1-1}, b_{n_1}, \dots, b_{n_2-1}, b_{n_1}, \dots, b_{n_2-1}, b_{n_2}, b_{n_2+1}, \dots, b_N, \dots] = [0; B_N^2, b_{N+1}, \dots].$$

Then $\max(\gamma^1, \gamma^2) > \gamma$.

Proof. We apply Lemma 4.6 for $A = (b_1, b_2, \dots, b_{n_1-1})$, $B = (b_{n_1}, b_{n_1+1}, \dots, b_{n_2-1})$, $C = (b_{n_2}, b_{n_2+1}, \dots)$. Here $\gamma = [0; A, B, C]$, $\gamma^1 = [0; A, C]$, and $\gamma^2 = [0; A, B, B, C]$. Note that as γ is not a quadratic irrationality, the sequence C is not periodic. □

Now we are ready to prove Theorem 1.

Proof. Suppose that B is not periodic on the right side. Consider an increasing sequence of indices $k(j)$ such that for any natural j the sequence $(a_{k(j)-j}, \dots, a_{k(j)}, \dots, a_{k(j)+j})$ coincides with the sequence $(b_{-j}, \dots, b_0, \dots, b_j)$. Of course,

$$\lim_{j\to\infty} \lambda_{k(j)}(\alpha) = \lambda_0(B) = \mu(\alpha).$$

Without loss of generality, one can say that $k(j+1) - k(j) \to \infty$ as $j \to \infty$. Consider an even n such that $\varepsilon_n < \frac{1}{2}(b-a)$ and $N = N(n)$ as defined in Lemma 4.5. Define $n_1 < n_2$ from Lemma 4.5 for the sequence $(b_1, \dots, b_N)$. As B is not periodic to the right, define a minimal positive integer r such that $b_{n_1+r} \neq b_{n_2+r}$. Consider the sequences B_N^1, B_N^2 and the continued fractions γ_1, γ_2 from Lemma 4.7 applied to the continued fraction $[0; b_1, \dots, b_n \dots] = \gamma$. If $\gamma_2 > \gamma$, define $g = 2$; otherwise we put $g = 1$. Consider the doubly infinite sequence $B' = (\dots, b_{-n}, b_0, B_N^g, b_{N+1}, \dots)$. Note that

$$a = \lambda_0(B) < \lambda_0(B') < a + \varepsilon_n < b.$$

Consider the corresponding continued fraction α' which is obtained from the continued fraction α by replacing every segment $(a_{k(j)}, \dots, a_{k(j)+N}) = (a_{k(j)}, B_N)$ by the segment $(a_{k(j)}, B_N^g)$ for every $j \geqslant n_2 + r$. One can easily see that α' and α satisfy the condition of Lemma 4.4 and hence $\mu(\alpha') < \mu(\alpha) + 2\varepsilon_n$. But as $\mu(\alpha) + 2\varepsilon_n < b$ and (a, b) is the gap in $\mathbb{L}$, we have

$$\mu(\alpha') \leqslant \mu(\alpha) = a. \tag{3}$$

On the other hand, one can easily see that the sequence B' is weakly associated with α'. This means that

$$\mu(\alpha') \geqslant M(B) \geqslant \lambda_0(B') > \lambda_0(B) = a.$$

We obtain a contradiction with (3). The case when B is not periodic on the left side is considered in exactly the same way. □

5. Proof of Theorem 2

The following lemma from [Gayfulin 2017] immediately implies the "⇐" part of the statement of Theorem 2.

Lemma 5.1. *Consider an arbitrary point a in the Lagrange spectrum. If there exists a quadratic irrationality γ such that $\mu(\gamma) = a$, then a is admissible.*

Now it is sufficient to prove that if a is an admissible left endpoint of a gap in the Lagrange spectrum, then there exists a quadratic irrationality α such that $\mu(\alpha) = a$.

Proof. Let a be an admissible left endpoint of some gap in the Lagrange spectrum. Let $\alpha = [a_0; a_1, \ldots, a_n, \ldots]$ be an irrational number such that $\mu(\alpha) = a$. Suppose that α is attainable, but not a quadratic irrationality. Let $k(j)$ be a growing sequence of indices such that

$$\lambda_{k(j)}(\alpha) \geqslant \mu(\alpha). \tag{4}$$

Of course,

$$\lim_{j\to\infty} \lambda_{k(j)}(\alpha) = \mu(\alpha).$$

Consider a sequence $B = (\ldots, b_{-n}, \ldots, b_{-1}, b_0, b_1, \ldots, b_n, \ldots)$ strongly associated with α having the following property: the sequence $(b_{-i}, \ldots, b_0, \ldots, b_i)$ coincides with the sequence $(a_{k(j)-i}, \ldots, a_{k(j)}, \ldots, a_{k(j)+i})$ for infinitely many j's. Theorem 1 implies that B is eventually periodic. That is, there exist a positive integer m and two finite sequences L and R such that

$$B = (\bar{L}, b_{-m}, \ldots, b_0, \ldots, b_m, \bar{R}).$$

It follows from (4) that one of the inequalities

$$[a_{k(j)}; a_{k(j+1)}, \ldots] \geqslant [b_0; b_1, \ldots, b_m, \bar{R}],$$
$$[0; a_{k(j-1)}, \ldots, a_1] \geqslant [0; b_{-1}, \ldots, b_{-m}, \bar{L}]$$

holds for infinitely many j's. Note that $[a_{k(j)}; a_{k(j+1)}, \ldots] \neq [b_0; b_1, \ldots, b_m, \bar{R}]$, as α is not a quadratic irrationality and, of course, $[0; a_{k(j-1)}, \ldots, a_1] \neq [0; b_{-1}, \ldots, b_{-m}, \bar{L}]$. Suppose that

$$[a_{k(j)}; a_{k(j+1)}, \ldots] > [b_0; b_1, \ldots, b_m, \bar{R}] \tag{5}$$

for infinitely many j's. Denote by p the length of period R. Denote by $r(j)$ the minimal positive number such that $a_{k(j)+r(j)} \neq b_{r(j)}$. Without loss of generality, one can say that:

(1) $k(j+1) - k(j) - r(j) \to \infty$ as $j \to \infty$.

(2) $[a_{k(j)}; a_{k(j+1)}, \ldots] > [b_0; b_1, \ldots, b_m, \bar{R}]$ for every $j \in \mathbb{N}$.

(3) $[a_{k(j)}; a_{k(j+1)}, \ldots, a_{k(j)+m}] = [b_0; b_1, \ldots, b_m]$ for every $j \in \mathbb{N}$.

(4) The sequence $(a_{k(j)-j}, \dots, a_{k(j)}, \dots, a_{k(j)+j})$ coincides with the sequence $(b_{-j}, \dots, b_0, \dots, b_j)$ for every $j \in \mathbb{N}$.

(5) Period length p is even.

Denote by $t(j)$ the number of periods P in the sequence $(b_{m+1}, \dots, b_{r(j)})$. Of course,

$$t(j) = \left[\frac{r(j)-m}{p}\right]$$

and $t(j)$ tends to infinity as $j \to \infty$. Lemma 4.3 implies that since (5) holds, we have

$$[a_{k(j)}; a_{k(j+1)}, \dots, a_{k(j)+m}, \underbrace{R, \dots, R}_{t(j) \text{ times}}, \dots, a_{k(j)+r(j)}, \dots] > [b_0; b_1, \dots, b_m, \bar{R}].$$

Denote by α_n a continued fraction obtained from the continued fraction $\alpha = [a_0; a_1, \dots, a_n, \dots]$ as follows: for any $j \in \mathbb{N}$ if $t(j) > n$, then every pattern

$$a_{k(j)}, a_{k(j+1)}, \dots, a_{k(j)+m}, \underbrace{R, \dots, R}_{t(j) \text{ times}}, \dots, a_{k(j)+r(j)}$$

is replaced by the pattern

$$a_{k(j)}, a_{k(j+1)}, \dots, a_{k(j)+m}, \underbrace{R, \dots, R}_{n \text{ times}}, \dots, a_{k(j)+r(j)}.$$

As the length of the period R is even, by Lemma 4.3 one has

$$[a_{k(j)}; a_{k(j+1)}, \dots, a_{k(j)+m}, \underbrace{R, \dots, R}_{n \text{ times}}, \dots, a_{k(j)+r(j)}, \dots] - [b_0; b_1, \dots, b_m, \bar{R}] > \delta_{m+(n+1)p}. \tag{6}$$

On the other hand, as the sequence $(a_{k(j)-j}, \dots, a_{k(j)})$ coincides with the sequence $(b_{-j}, \dots, b_0)$ for all $j \in \mathbb{N}$, by Lemma 4.3 one has

$$\left|[0; a_{k(j)-1}, \dots, a_{k(j)-j}, \dots, a_1] - [0; b_{-1}, \dots, b_{-m}, \bar{L}]\right| < \varepsilon_j. \tag{7}$$

For any positive integers n, m, p there exists J such that for all $j > J$ one has $\varepsilon_j < \frac{1}{2}\delta_{m+(n+1)p}$. Now, from (6) and (7) we have for $j > J$

$$\begin{aligned}\big([0; a_{k(j)-1}, \dots, a_{k(j)-j}, \dots, a_1] \\ + [a_{k(j)}; a_{k(j+1)}, \dots, a_{k(j)+m}, \underbrace{R, \dots, R}_{n \text{ times}}, \dots, a_{k(j)+r(j)}, \dots]\big) \\ - \big([0; b_{-1}, \dots, b_{-m}, \bar{L}] + [b_0; b_1, \dots, b_m, \bar{R}]\big) > \tfrac{1}{2}\delta_{m+(n+1)p}.\end{aligned} \tag{8}$$

Considering the limit in (8) as $j \to \infty$, we can easily see that $\mu(\alpha_n) \geqslant \mu(\alpha) + \frac{1}{2}\delta_{m+(n+1)p}$.

Note that

$$\lim_{n\to\infty} \mu(\alpha_n) = \mu(\alpha) = a. \tag{9}$$

Indeed, every pattern of length np which occurs in the sequence of partial quotients of α infinitely many times occurs in the sequence of partial quotients of α_n infinitely many times. Similarly, every pattern of

length np which occurs in the sequence of partial quotients of α_n infinitely many times, occurs in the sequence of partial quotients of α infinitely many times too. Then, by Lemma 4.4,

$$|\mu(\alpha) - \mu(\alpha_n)| < 2\varepsilon_{np} = 2^{-np+2} \to 0$$

as $n \to \infty$. We obtain a contradiction with the fact that a is the left endpoint of the gap (a, b) in the Lagrange spectrum. Indeed, (8) implies that $\mu(\alpha_n) > \mu(\alpha)$ for all $n \in \mathbb{N}$. In addition, (9) implies that there exists a positive integer N such that for any $n > N$ one has $a = \mu(\alpha) < \mu(\alpha_n) < b$.

If (5) does not hold infinitely many times, then the inequality

$$[0; a_{k(j-1)}, \dots, a_1] > [0; b_{-1}, \dots, b_{-m}, \bar{L}]$$

holds infinitely many times. As α is not a quadratic irrationality, for any positive integer s there exists an integer $N(s) > s$ such that for all $n \geqslant N(s)$ the continued fraction $[0; a_n, a_{n-1}, \dots, a_s]$ is not convergent to the continued fraction $[0; b_{-1}, \dots, b_{-m}, \bar{L}]$. Without loss of generality, one can say that $k(1) > N(1)$, $k(j+1) > N(2k(j))$. Denote by $r(j)$ the minimal positive number such that $a_{k(j)-r(j)} \neq b_{-r(j)}$. It is easy to see that the number $r(j)$ is well-defined and

$$r(j+1) \leqslant k(j+1) - N(2k(j)) < k(j+1) - 2k(j).$$

Therefore $k(j+1) - r(j+1) - k(j) \to \infty$ as $j \to \infty$. Now one can easily complete the proof using exactly the same argument as we used in the first case. □

6. Proof of Theorem 3

First of all, let us show that (α_n^*, β_n) is the maximal gap in $\mathbb{L}$. As

$$\beta_n = 2 + 2[0; \underbrace{\overline{1, \dots, 1}}_{2n}, 2] = \mu([0; \underbrace{\overline{1, \dots, 1}}_{2n}, 2]),$$

we have $\beta_n \in \mathbb{L}$. The proof of the fact that $\alpha_n^* \in \mathbb{L}$ when $n \geqslant 2$ is a little more complicated. Recall that

$$\alpha_n^* = 2 + [0; \underbrace{1, \dots, 1}_{2n-2}, \overline{2, 2, 1, 2}] + [0; \underbrace{1, \dots, 1}_{2n-1}, 2, \underbrace{1, \dots, 1}_{2n-2}, \overline{2, 2, 1, 2}].$$

Denote by $C_n(k)$ the finite sequence of integers

$$C_n(k) = (\underbrace{2, 1, 2, 2}_{k}, \underbrace{1, \dots, 1}_{2n-2}, 2, \underbrace{1, \dots, 1}_{2n-1}, 2, \underbrace{1, \dots, 1}_{2n-2}, \underbrace{2, 2, 1, 2}_{k}).$$

A little calculation shows that

$$L(\overline{C_n(k)}) = 2 + [0; \underbrace{1, \dots, 1}_{2n-2}, \underbrace{2, 2, 1, 2}_{k}, \dots] + [0; \underbrace{1, \dots, 1}_{2n-1}, 2, \underbrace{1, \dots, 1}_{2n-2}, \underbrace{2, 2, 1, 2}_{k}, \dots].$$

Therefore $\lim_{k\to\infty} L(\overline{C_n(k)}) = \alpha_n^*$. As $\mathbb{L}$ is closed set, we obtain that $\alpha_n^* \in \mathbb{L}$.

By [Gbur 1976, Lemma 4], α_n^* is a growing sequence. One can easily see that

$$\lim_{n\to\infty} \alpha_n^* = 2 + 2[0; \bar{1}] = \sqrt{5} + 1 \approx 3.236.$$

Thus, we have

$$\alpha_2^* \leqslant \alpha_n^* < 1+\sqrt{5}, \quad \text{where } n \geqslant 2.$$

The following lemma is a compilation of Lemmas 3 and 4 from [Gbur 1976].

Lemma 6.1. *Consider a doubly infinite sequence* $B = (\ldots, b_{-n}, \ldots, b_{-1}, b_0, b_1, \ldots, b_n, \ldots)$ *such that* $M(B) < \sqrt{5}+1$. *Then all elements of* B *are bounded by* 2 *and* B *does not contain patterns of the form* $(2,1,2,1)$ *and* $(1,2,1,2)$.

By Lemma 4.1, without loss of generality one can say that $M(B) = \lambda_0(B)$. Denote the continued fractions $[0; b_1, \ldots, b_n, \ldots]$ and $[0; b_{-1}, \ldots, b_{-n}, \ldots]$ by x and y respectively. Then

$$M(B) = b_0 + x + y.$$

Without loss of generality one can say that $x \leqslant y$. Now we need the following lemma from [Gbur 1976, Theorem 4(i)].

Lemma 6.2. *Let* B *be a doubly infinite sequence such that* $M(B) = \lambda_0(B)$. *Then for all* $n \geqslant 1$ *we have*

$$\beta_n \leqslant M(B) = 2 + x + y \leqslant \alpha_{n+1}^* \quad \Longleftrightarrow \quad x = [0; \underbrace{1, \ldots, 1}_{2n}, 2, \ldots] \text{ and } y = [0; \underbrace{1, \ldots, 1}_{2n}, \ldots].$$

It also follows from [Gbur 1976, Theorem 4(ii)] that

$$2 + [0; \underbrace{1, \ldots, 1}_{2n+1}, \ldots] + [0; \underbrace{1, \ldots, 1}_{2n+1}, 2, \ldots] < \sqrt{5}+1.$$

Define

$$w_0 = [0; \overline{2, 1, 2, 2}],$$

$$x_0 = [0; \underbrace{1, \ldots, 1}_{2n}, \overline{2, 2, 1, 2}] = [0; \underbrace{1, \ldots, 1}_{2n}, 2 + w_0],$$

$$y_0 = [0; \underbrace{1, \ldots, 1}_{2n+1}, 2, \underbrace{1, \ldots, 1}_{2n}, \overline{2, 2, 1, 2}] = [0; \underbrace{1, \ldots, 1}_{2n+1}, 2, \underbrace{1, \ldots, 1}_{2n}, 2 + w_0].$$

Lemma 6.3. *Let* $w = [0; a_1, a_2, \ldots, a_n, \ldots]$ *be a continued fraction with elements equal to* 1 *or* 2. *Suppose that the sequence* $(a_1, a_2, \ldots, a_n, \ldots)$ *does not contain the pattern* $(2,1,2,1)$. *Then* $w \geqslant w_0$.

Proof. Denote the elements of the continued fraction $w_0 = [0; \overline{2, 1, 2, 2}]$ by $[0; a_1', \ldots, a_m', \ldots]$. Denote by r the minimal index such that $a_r \neq a_r'$. Suppose that $w < w_0$. Then either r is odd, $a_r = 2,\ a_r' = 1$ or r is even, $a_r = 1,\ a_r' = 2$. However $a_r' = 2$ for any odd r; thus the first case leads to a contradiction. Consider the second case. Of course, $r \geqslant 4$. Then $a_{r-3}' = a_{r-3} = 2,\ a_{r-2}' = a_{r-2} = 1,\ a_{r-1}' = a_{r-1} = 2$. This means that $(a_{r-3}, a_{r-2}, a_{r-1}, a_r) = (2,1,2,1)$ and we obtain a contradiction. □

Now we prove Theorem 3.

Proof. Suppose that α_n^* is admissible for some $n \geqslant 2$. Consider an attainable number α such that $\mu(\alpha) = \alpha_n^*$. Let $B = (\ldots, b_{-n}, \ldots, b_{-1}, b_0, b_1, \ldots, b_n, \ldots)$ be any sequence strongly associated with α. Denote by f the increasing function

$$f(t) = [0; \underbrace{1, \ldots, 1}_{2n+1}, 2 + t].$$

By Lemma 6.2, there exist $0 < v, w < 1$ such that

$$x = [0; \underbrace{1, \dots, 1}_{2n}, 2 + v] \quad \text{and} \quad y = [0; \underbrace{1, \dots, 1}_{2n-1}, 1 + w].$$

Note that $x \leqslant x_0$. Indeed,

$$x = [0; \underbrace{1, \dots, 1}_{2n}, 2 + v] \leqslant [0; \underbrace{1, \dots, 1}_{2n}, 2 + w_0] = x_0 \quad \Longleftrightarrow \quad v \geqslant w_0.$$

The last equality follows from Lemmas 6.3 and 6.1. Therefore

$$y = [0; \underbrace{1, \dots, 1}_{2n-1}, 1 + w] \geqslant y_0 = f(x_0).$$

In particular, $b_{-2n-1} = 1$, $b_{-2n-2} = 2$ and there exists $0 < v \leqslant x_0$ such that $y = f(v) \geqslant f(x_0)$. Hence $v \geqslant x_0$. On the other hand,

$$2 + v + f(x) = \lambda_{-2n-2}(B) \leqslant M(B) = \lambda_0(B) = 2 + x + f(v).$$

As $|f(y) - f(z)| < |y - z|$ for any $0 < y, z < 1$, one can easily see that $v \leqslant x$. Thus, $v = x = x_0$ and $y = y_0$. Hence the sequence B satisfies

$$B = (\overline{2, 1, 2, 2}, \underbrace{1, \dots, 1}_{2n}, 2, \underbrace{1, \dots, 1}_{2n+1}, 2, \underbrace{1, \dots, 1}_{2n}, \overline{2, 2, 1, 2})$$

and is not purely periodic. We obtain a contradiction with Corollary 3.1, as we supposed B to be an arbitrary sequence strongly associated with α. □

References

[Cusick 1975] T. W. Cusick, "The connection between the Lagrange and Markoff spectra", *Duke Math. J.* **42**:3 (1975), 507–517. MR Zbl

[Cusick and Flahive 1989] T. W. Cusick and M. E. Flahive, *The Markoff and Lagrange spectra*, Mathematical Surveys and Monographs **30**, Amer. Math. Soc., Providence, RI, 1989. MR Zbl

[Freiman 1973] G. A. Freiman, "The initial point of Hall's ray", pp. 87–125 in *Number-theoretic studies in the Markov spectrum and in the structural theory of set addition*, edited by G. A. Freiman et al., Kalinin. Gosudarstv. Univ., Moscow, 1973. In Russian. MR Zbl

[Gayfulin 2017] D. Gayfulin, "Attainable numbers and the Lagrange spectrum", *Acta Arith.* **179**:2 (2017), 185–199. MR Zbl

[Gbur 1976] M. E. Gbur, "On the lower Markov spectrum", *Monatsh. Math.* **81**:2 (1976), 95–107. MR Zbl

[Malyshev 1977] A. V. Malyshev, "Markov and Lagrange spectra (a survey of the literature)", *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov.* **67** (1977), 5–38. In Russian; translated in *J. Sov. Math.* **16**:1 (1981), 767–788. MR Zbl

[Schecker 1977] H. Schecker, "Über die Menge der Zahlen, die als Minima quadratischer Formen auftreten", *J. Number Theory* **9**:1 (1977), 121–141. MR Zbl

Received 10 Jan 2018.

DMITRY GAYFULIN:

gayfulin@rambler.ru

Steklov Mathematical Institute, Russian Academy of Sciences, Moscow, Russia

msp

Moscow Journal of Combinatorics and Number Theory
Vol. 8, No. 1, 2019
dx.doi.org/10.2140/moscow.2019.8.57

Transcendence of numbers related with Cahen's constant

Daniel Duverney, Takeshi Kurosawa and Iekata Shiokawa

Cahen's constant is defined by the alternating sum of reciprocals of terms of Sylvester's sequence minus 1. Davison and Shallit proved the transcendence of the constant and Becker improved it. In this paper, we study rationality of functions satisfying certain functional equations and generalize the result of Becker by a variant of Mahler's method.

1. Introduction

Sylvester's sequence $\{S_n\}_{n\geq 0}$ is defined by the recurrence

$$S_0 = 2, \qquad S_{n+1} = S_n^2 - S_n + 1 \quad (n \geq 0).$$

It is well known that

$$\sum_{n=0}^{\infty} \frac{1}{S_n} = 1. \tag{1}$$

Cahen [1891] showed that the number

$$C = \sum_{n=0}^{\infty} \frac{(-1)^n}{S_n - 1}, \tag{2}$$

which is now called Cahen's constant, is irrational. Davison and Shallit [1991] established the transcendence of Cahen's constant. They constructed a class of alternating series each of which can be expanded in an explicit simple continued fraction having irrationality exponent greater than 2.5 and showed that the series (2) belongs to this class. Here, for an irrational number α, the irrationality exponent $\mu(\alpha)$ is defined by the least upper bound of the set of numbers μ for which the inequality

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{\mu}}$$

has infinitely many irreducible rational solutions p/q. Thus, the transcendence of Cahen's constant C follows from Roth's theorem. Becker [1992, Corollary 3] improved the result by a variant of Mahler's method. Indeed, he proved the following: Let $p(z)$ be a polynomial with algebraic coefficients and $\deg p(z) \geq 2$ and $q(z) = z - \gamma$ with an algebraic number γ. Let x be an algebraic number such that

MSC2010: 11J81.
Keywords: Cahen's constant, transcendence, Mahler's method, Sylvester's sequence.

$\lim_{n\to\infty} p^n(x) = \infty$ and $q(p^n(x)) \neq 0$ for all $n \geq 0$, where $p^0(z) = z$, $p^n(z) = p(p^{n-1}(z))$ $(n \geq 1)$. Then, the number

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{q(p^n(x))}$$

is transcendental except when $q(p(z)) = \lambda^{-1}q(z)^2 + q(z) - \lambda$ for some constant $\lambda \neq 0$, in which case

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{q(p^n(z))} = \frac{1}{q(z)+\lambda}.$$

For example, if $p(z) = z^2 - z + 1$ and $\alpha = S_0$, then the number

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{S_n - \gamma}$$

is transcendental for any algebraic γ with $S_n \neq \gamma$ for all $n \geq 0$.

In this paper, we consider the function

$$f(z) = \sum_{n=0}^{\infty} \frac{a^n}{q(p^n(z))}, \tag{3}$$

where $a \neq 0$ is a complex number, $p(z) \in \mathbb{C}[z]$ with $\deg p(z) \geq 2$, and $q(z) \in \mathbb{C}[z]$ with $\deg q(z) \geq 1$. We note that the right-hand side of (3) is convergent at any $z \in \mathbb{C}$ for which $\lim_{n\to\infty} p^n(z) = \infty$ and $q(p^n(z)) \neq 0$ for all $n \geq 0$. Furthermore, there exists a constant $C_f > 1$ such that $f(z)$ is analytic in $\mathcal{D}_f = \{z \in \mathbb{C} \mid |z| > C_f\}$ and $f(\mathcal{D}_f) \subset \mathcal{D}_f$.

The function $f(z)$ satisfies the functional equation

$$af(p(z)) = f(z) - \frac{1}{q(z)}, \tag{4}$$

and more generally

$$f(p^n(z)) = \frac{1}{a^n}\left(f(z) - \sum_{j=0}^{n-1} \frac{a^j}{q(p^j(z))}\right) \quad (n \geq 1). \tag{5}$$

We now state our results.

Theorem 1.1. *Let $f(z)$ be the function defined by*

$$f(z) = \sum_{n=0}^{\infty} \frac{a^n}{q(p^n(z))},$$

where $a \in \mathbb{C}^\times$, $p(z) \in \mathbb{C}[z]$ with $\deg p(z) \geq 2$ and $q(z) \in \mathbb{C}[z]$ is monic with $\deg q(z) \geq 1$. Then, the function $f(z)$ is algebraic over the field $\mathbb{C}(z)$ of rational functions if and only if $\deg p(z) = 2$ and $p(z)$ and $q(z)$ satisfy the relation

$$b^l q(p(z)) - a = b^l q(z)(b^l q(z) - a), \tag{6}$$

where b is the leading coefficient of $p(z)$ and $l = \deg q(z)$, and if so

$$f(z) = \frac{b^l}{b^l q(z) - a}. \tag{7}$$

Theorem 1.2. *With the same notation as in Theorem 1.1, assume that a and the coefficients of $p(z)$ and $q(z)$ are algebraic. Then the number*

$$f(x) = \sum_{n=0}^{\infty} \frac{a^n}{q(p^n(x))}$$

is transcendental for any algebraic x with $\lim_{n\to\infty} p^n(x) = \infty$ and $q(p^n(x)) \neq 0$ for all $n \geq 0$, except when $d = 2$ and $p(z)$ and $q(z)$ satisfy the relation (6)*, in which case $f(z)$ is the rational function given by* (7)*.*

Theorem 1.3. *Let $f(z)$ be the function defined by*

$$f(z) = \sum_{n=0}^{\infty} \frac{a^n}{(p^n(z) - \gamma)^l}, \tag{8}$$

where $p(z) \in \mathbb{C}[z]$ with $\deg p(z) \geq 2$ and l is a positive integer. Assume that $a \neq 0$, γ, and the coefficients of $p(z)$ are algebraic numbers. Then the value $f(x)$ is transcendental for any algebraic x with $\lim_{n\to\infty} p^n(x) = \infty$ and $p^n(x) \neq \gamma$ for all $n \geq 0$, except in the following two cases:

(i) $l = 1$, $p(\gamma) - \gamma + b^{-1}p'(\gamma) = 0$ *and* $a = -p'(\gamma)$, *in which case*

$$f(x) = \frac{b}{b(x-\gamma) - a}. \tag{9}$$

(ii) $l = 2$, $p(\gamma) - \gamma = -2b^{-1}$, $p'(\gamma) = 0$ *and* $a = 4$, *in which case*

$$f(x) = \frac{b^2}{b^2(x-\gamma)^2 - 4}. \tag{10}$$

Remark 1.4. The case (ii) can be obtained as a special case of (i). Indeed, if $a = 4$ in case (ii), we have by Taylor's formula $p(z) = b(x-\gamma)^2 - 4(x-\gamma) + \gamma + 4b^{-1}$, and therefore $p(z) - \gamma = b(x - \gamma - 2b^{-1})^2$. Hence

$$f(x) = \frac{1}{x-\gamma} + \sum_{n=1}^{\infty} \frac{4^n}{p(p^{n-1}(x)) - \gamma} = \frac{1}{x-\gamma} + 4b^{-1} \sum_{n=0}^{\infty} \frac{4^n}{(p^n(x) - \gamma - 2b^{-1})^2}.$$

Replacing $f(x)$ by using (9) and $\gamma + 2b^{-1}$ by γ yields

$$\frac{b}{b(x-\gamma) - 2} = \frac{b}{b(x-\gamma) + 2} + 4b^{-1} \sum_{n=0}^{\infty} \frac{4^n}{(p^n(x) - \gamma)^2},$$

which is exactly (10).

We give some examples of Theorem 1.3.

Example 1.5. Let $\{S_n\}_{n\geq 0}$ be Sylvester's sequence defined by

$$S_{n+1} = S_n^2 - S_n + 1 \quad (n \geq 0)$$

with arbitrary $S_0 \in \mathbb{Z} \setminus \{0, 1\}$. Here $p(z) = z^2 - z + 1$, $p'(z) = 2z - 1$ and $b = 1$. Let us study first case (i) in Theorem 1.3. The equation $p(\gamma) - \gamma + b^{-1}p'(\gamma) = 0$ is equivalent to $\gamma^2 = 0$. Therefore $\gamma = 0$ and $a = -p'(\gamma) = 1$. Case (ii) cannot occur. Hence for any algebraic numbers $a \neq 0$ and γ with $S_n \neq \gamma$ for all $n \geq 0$ and a positive integer l, the number

$$\sum_{n=0}^{\infty} \frac{a^n}{(S_n - \gamma)^l}$$

is transcendental except when $l = a = 1$ and $\gamma = 0$, and if so

$$\sum_{n=0}^{\infty} \frac{1}{S_n} = \frac{1}{S_0 - 1}.$$

Example 1.6. Let $\{T_n\}_{n\geq 0}$ be the recurrence

$$T_0 \in \mathbb{Z}, \quad |T_0| > 2, \qquad T_{n+1} = T_n^2 - 2 \quad (n \geq 0).$$

Here $p(z) = z^2 - 2$, $p'(z) = 2z$ and $b = 1$. By Theorem 1.3, we see that, for any algebraic numbers $a \neq 0$ and γ with $T_n \neq \gamma$ for all $n \geq 0$ and a positive integer l, the number

$$\sum_{n=0}^{\infty} \frac{a^n}{(T_n - \gamma)^l}$$

is transcendental except in the following three cases:

(i) $l = 1$, $\gamma = 1$, and $a = -2$, in which case

$$\sum_{n=0}^{\infty} \frac{(-2)^n}{T_n - 1} = \frac{1}{T_0 + 1}. \tag{11}$$

(ii) $l = 1$, $\gamma = -2$, and $a = 4$, in which case

$$\sum_{n=0}^{\infty} \frac{4^n}{T_n + 2} = \frac{1}{T_0 - 2}. \tag{12}$$

(iii) $l = 2$, $\gamma = 0$, and $a = 4$, in which case

$$\sum_{n=0}^{\infty} \frac{4^n}{T_n^2} = \frac{1}{(T_0 - 2)(T_0 + 2)}.$$

As mentioned in Remark 1.4, (iii) is intrinsically the same as (ii).

Example 1.7. Fermat numbers $F_n = 2^{2^n} + 1$ satisfy the recurrence relation

$$F_{n+1} = F_n^2 - 2F_n + 2 \quad (n \geq 0)$$

with $F_0 = 3$. By Theorem 1.3, for any algebraic numbers $a \neq 0$ and γ with $F_n \neq \gamma$ for all $n \geq 0$ and a positive integer l, the number

$$\sum_{n=0}^{\infty} \frac{a^n}{(F_n - \gamma)^l} \tag{13}$$

is transcendental except when $l = 1$, $a = 2$, and $\gamma = 0$, and if so

$$\sum_{n=0}^{\infty} \frac{2^n}{F_n} = \frac{1}{F_0 - 2} = 1. \tag{14}$$

Remark 1.8. Formulas (11), (12), and (14) are known; see formulas (2.22), (2.25), and (2.26) in [Duverney 2001]. In fact, let α and β with $|\alpha| > |\beta|$ be roots of the equations $x^2 - T_0 x - 1 = 0$. Then the Lucas-type sequence

$$T_n = \alpha^{2^n} + \beta^{2^n}$$

satisfies $T_{n+1} = T_n^2 - 2$ $(n \geq 1)$. Therefore the series (11) and (12), as well as (14), can also be seen as examples of exceptional cases related to the classical Mahler's method; see [Duverney et al. 2002, Theorem 1.3; Kanoko et al. 2009, Example 1].

2. Proof of Theorems 1.1 and 1.3

To prove the theorems, we study rational solutions of a functional equation which generalizes (4).

Lemma 2.1. *Let $a, c \in \mathbb{C}^\times$, $p(z) \in \mathbb{C}[z]$ with $d = \deg p(z) \geq 2$ and leading coefficient b, and $q(z) \in \mathbb{C}[z]$ be monic with $l = \deg q(z) \geq 1$. Assume that a rational function $g(z)$ satisfies the functional equation*

$$ag(p(z)) = g(z) - \frac{\delta}{q(z)}. \tag{15}$$

Then $d = 2$, and $p(z)$ and $q(z)$ satisfy the relation

$$b^l q(p(z)) - a = b^l q(z)(b^l q(z) - a), \tag{16}$$

in which case:

(i) *If $a \neq 1$, then* (15) *has one and only one rational solution, which is*

$$g(z) = \frac{\delta}{q(z) - ab^{-l}}. \tag{17}$$

(ii) *If $a = 1$, then* (15) *has infinitely many rational solutions given by*

$$g(z) = \alpha + \frac{\delta}{q(z) - b^{-1}} \quad (\alpha \in \mathbb{C}). \tag{18}$$

Proof. Let $R(z)$ and $S(z)$ be two coprime monic polynomials and $\alpha \in \mathbb{C}^\times$ be such that

$$g(z) = \alpha \frac{R(z)}{S(z)}.$$

As $g(z)$ satisfies (15), we have for $c = \delta\alpha^{-1}$

$$a\frac{R(p(z))}{S(p(z))} = \frac{R(z)}{S(z)} - \frac{c}{q(z)}. \tag{19}$$

Put for brevity $r = \deg R(z)$ and $s = \deg S(z)$. If $s = 0$, then there is no solution satisfying (19) since $g(z) - ag(p(z)) = c/(q(z)) \notin \mathbb{C}[z]$. Hence, $s \geq 1$. The functional equation (19) can be written as

$$aR(p(z))S(z)q(z) = R(z)S(p(z))q(z) - cS(z)S(p(z)). \tag{20}$$

Since $(R(p(z)), S(p(z))) = 1$, we have

$$S(p(z)) \mid S(z)q(z). \tag{21}$$

Hence, $ds \leq s + l$. Therefore, we obtain

$$1 \leq s \leq \frac{l}{d-1}. \tag{22}$$

Comparing the degrees of both sides of (20), we get $r \leq s$.

If $r < s$, the degree of the first term of the right-hand side in (20) is greater than that of the left-hand side. Therefore, the degree of the first term of the right-hand side is equal to that of the second term of the right-hand side. Then, we have using (22)

$$0 = r + ds + l - (s + ds) \geq r - s + (d-1)s = r + (d-2)s \geq 0.$$

Therefore, we deduce $d = 2$ and $r = 0$. This together with (20) leads to

$$aS(z)q(z) = S(p(z))q(z) - cS(z)S(p(z)). \tag{23}$$

The degree of the left-hand side is less than that of the first term of the right-hand side. Hence, the degrees of the two terms in the right-hand side are equal, and so $s = l$. This and (21) with $d = 2$ imply

$$S(p(z)) = b^l q(z)S(z). \tag{24}$$

Substituting (24) in (23), we get $a = b^l(q(z) - cS(z))$. Comparing the leading coefficients of both sides, we find $c = 1$ and

$$S(z) = q(z) - ab^{-l}. \tag{25}$$

Substituting into (24) yields (16). In this case, as $R(z)$ is monic and $\deg R(z) = 0$, we have $R(z) = 1$ and

$$g(z) = \alpha\frac{R(z)}{S(z)} = c^{-1}\delta\frac{1}{q(z) - ab^{-l}},$$

which proves that (17) holds (also for $a = 1$).

Now, let $r = s$. Then we get $a = 1$ by comparing the leading coefficients of both sides in (20). Put $T(z) = R(z) - S(z)$. Then, by (20),

$$T(p(z))S(z)q(z) = T(z)S(p(z))q(z) - cS(z)S(p(z)). \tag{26}$$

Noting that $\deg T(z) < s$ and $(S(z), T(z)) = 1$, we apply the above discussion for $S(z)$ and $T(z)$, and thus we obtain $d = 2$, $T(z)$ is a constant, and (24). Let $T(z) = k \neq 0$. Substituting (24) into (26), we get

$$1 = b^l(q(z) - ck^{-1}S(z)).$$

Comparing the leading coefficients of both sides, we find $k = c$, and we see that (25) holds again. Therefore (16) holds. In this case $R(z) - S(z) = c$, whence

$$g(z) = \alpha \frac{R(z)}{S(z)} = \alpha + \frac{\alpha c}{q(z) - b^{-l}},$$

which proves (18). □

Now we prove Theorem 1.1 by using Lemma 2.1.

Proof of Theorem 1.1. Assume that the function (3) is algebraic over the field $\mathbb{C}(z)$ of rational functions. Then we have

$$(f(z))^{\delta} + g(z)(f(z))^{\delta-1} + \cdots = 0, \tag{27}$$

where the degree δ is chosen to be minimal and $g(z)$ is a rational function with complex coefficients. Replacing z by $p(z)$ in (27) yields

$$\left(\frac{1}{a}\left(f(z) - \frac{1}{q(z)}\right)\right)^{\delta} + g(p(z))\left(\frac{1}{a}\left(f(z) - \frac{1}{q(z)}\right)\right)^{\delta-1} + \cdots = 0$$

by using (4). This can be written as

$$f(z)^{\delta} + \left(ag(p(z)) - \frac{\delta}{q(z)}\right) f(z)^{\delta-1} + \cdots = 0. \tag{28}$$

As δ is minimal, comparison with (27) and (28) yields

$$ag(p(z)) = g(z) + \frac{\delta}{q(z)}.$$

Since $g(z)$ satisfies the functional equation (15), we can apply Lemma 2.1 and obtain (6). Replacing z by $p^n(z)$ in (6) yields

$$\frac{ab^l}{b^l q(p^{n+1}(z)) - a} = \frac{a}{q(p^n(z))(b^l q(p^n(z)) - a)} = \frac{b^l}{b^l q(p^n(z)) - a} - \frac{1}{q(p^n(z))}.$$

After multiplying by a^n, the function $f(z)$ appears as a telescoping series and we have

$$f(z) = b^l \sum_{n=0}^{\infty} \left(\frac{a^n}{b^l q(p^n(z)) - a} - \frac{a^{n+1}}{b^l q(p^{n+1}(z)) - a} \right) = \frac{b^l}{b^l q(z) - a}.$$ □

Lemma 2.2. *Make the same assumptions as in Lemma 2.1. Let $q(z) = (z - \gamma)^l$, where $l \geq 1$. Then $l = 1$ or 2:*

(i) *If $l = 1$, then $b(p(\gamma) - \gamma) + p'(\gamma) = 0$ and $a = -p'(\gamma)$.*

(ii) *If $l = 2$, then $p'(\gamma) = 0$, $p(\gamma) - \gamma = -2b^{-1}$, and $a = 4$.*

Proof. By Lemma 2.1, (16) holds and we get

$$(p(z) - \gamma)^l - ab^{-l} = b^l (z - \gamma)^{2l} - a(z - \gamma)^l. \tag{29}$$

Differentiating both sides of (29), we get

$$p'(z)(p(z)-\gamma)^{l-1} = 2b^l(z-\gamma)^{2l-1} - a(z-\gamma)^{l-1}. \tag{30}$$

If $l = 1$, then taking $z = \gamma$ yields $p(z) - \gamma - ab^{-l} = 0$ and $p'(\gamma) = -a$. Replacing a in the first equality gives $b(p(\gamma) - \gamma) + p'(\gamma) = 0$, as claimed.

Let $l \geq 2$. By (29), we have

$$(p(\gamma)-\gamma)^l = ab^{-l} \neq 0. \tag{31}$$

Since $p(\gamma) \neq \gamma$ by (31), $(z-\gamma)^{l-1}$ divides $p'(z)$. Hence $l = 2$, and so (30) is reduced to

$$p(z) - \gamma = b(z-\gamma)^2 - \tfrac{1}{2}ab^{-1}. \tag{32}$$

Substituting $z = \gamma$ in (32) and using (31), we find $a = 4$ and $p(\gamma) - \gamma = -2b^{-1}$. Substituting $z = \gamma$ in (30) and using (31), we obtain $p'(\gamma) = 0$. □

Finally, we prove Theorem 1.3 by using Lemma 2.2 and Theorem 1.2, which will be shown independently in the next section using Theorem 1.1.

Proof of Theorem 1.3. If the function $f(z)$ defined in Theorem 1.3 is not a rational function, then the value $f(x)$ is transcendental by Theorem 1.2. Assume to the contrary that $f(z)$ is a rational function. Then Lemma 2.2 with (7) yields the exceptional cases. □

3. Proof of Theorem 1.2

Becker's result mentioned in Section 1 is a special case of the main theorem in [Becker 1992], which establishes algebraic independence of the values of power series $f_1(z), \ldots, f_m(z)$ satisfying the functional equations

$$f_i(z) = a_i(z) f_i(Tz) + b_i(z) \quad (i = 1, \ldots, m),$$

where $a_i(z), b_i(z)$ are rational functions with algebraic coefficients and $Tz = p(z^{-1})^{-1}$ for a polynomial $p(z)$ with algebraic coefficients and $\deg p(z) \geq 2$. The proof of this theorem is based on a deep result due to Philippon [1986] on a criterion for algebraic independence of complex numbers and is rather involved. Although Theorem 1.2 can also be deduced from [Becker 1992, Theorem], we give here a self-contained proof for completeness.

We prove Theorem 1.2 by a variant of Mahler's method. In the proof we will have to estimate the denominators and houses of algebraic numbers. We will use the following lemmas.

Lemma 3.1. *Let $\mathbb{K}$ be any algebraic field of degree k, and let $h \in \mathbb{K}[z]$. Let $\delta = \deg h$. Then there exists $\mu = \mu(h) \geq 1$ such that, for every $\theta \in \mathbb{K}^\times$,*

(i) $\operatorname{den} h(\theta) \leq \mu(\operatorname{den}\theta)^\delta$,

(ii) $\overline{|h(\theta)|} \leq \mu(\max(1, \overline{|\theta|}))^\delta$.

Proof. Put $h(z) = \sum_{i=0}^{\delta} a_i z^i$, with $a_\delta \neq 0$. Then clearly

$$\operatorname{den} h(\theta) \leq D(\operatorname{den}\theta)^\delta,$$

where $D = \mathrm{LCM}(\mathrm{den}\, a_1, \mathrm{den}\, a_2, \ldots, \mathrm{den}\, a_\delta)$. Moreover, denote by $\sigma_1 = \mathrm{Id}, \sigma_2, \ldots, \sigma_k$ the monomorphisms of $\mathbb{K}$. Then for every $j = 1, 2, \ldots, k$, we have

$$\left|\sum_{i=0}^{\delta} \sigma_j(a_i)(\sigma_j(\theta))^i\right| \le \sum_{i=0}^{\delta} |\sigma_j(a_i)|(\overline{|\theta|})^i \le \left(\sum_{i=0}^{\delta} |\sigma_j(a_i)|\right)(\max(1, \overline{|\theta|}))^\delta, \qquad \square$$

Lemma 3.2. *Let $\mathbb{K}$ be any algebraic field of degree k, and let $h \in \mathbb{K}[z]$. Let $\delta = \deg h$. Then for every $\theta \in \mathbb{K}^\times$ such that $h(\theta) \neq 0$, there exist $\nu = \nu(h) \ge 1$ such that*

$$\max\left(\mathrm{den}\left(\frac{1}{h(\theta)}\right), \overline{\left|\frac{1}{h(\theta)}\right|}\right) \le \nu(\mathrm{den}\,\theta \times \max(1, \overline{|\theta|}))^{k\delta}.$$

Proof. First we have

$$\mathrm{den}\left(\frac{1}{h(\theta)}\right) = \mathrm{den}\left(\frac{\mathrm{den}\, h(\theta)}{\mathrm{den}\, h(\theta) \times h(\theta)}\right) = \mathrm{den}\left(\frac{\mathrm{den}\, h(\theta) \prod_{i\neq 1} \sigma_i(\mathrm{den}\, h(\theta) \times h(\theta))}{N(\mathrm{den}\, h(\theta) \times h(\theta))}\right),$$

where $N(\alpha)$ is the norm of $\alpha \in \mathbb{K}$ over $\mathbb{Q}$. The numerator of the fraction is an integer of $\mathbb{K}$, and therefore

$$\mathrm{den}\left(\frac{1}{h(\theta)}\right) \le |N(\mathrm{den}\, h(\theta) \times h(\theta))| \le (\mathrm{den}\, h(\theta))^k \times \overline{|h(\theta)|}^k,$$

which proves the first part of Lemma 3.2 by using Lemma 3.1(i).

For the second part, for every $i = 1, 2, \ldots, k$, we have

$$\left|\sigma_i\left(\frac{1}{h(\theta)}\right)\right| = \left|\frac{\mathrm{den}\, h(\theta)}{\mathrm{den}\, h(\theta) \times \sigma_i(h(\theta))}\right| = \left|\frac{(\mathrm{den}\, h(\theta))^k \times \prod_{j\neq i} \sigma_j(h(\theta))}{N(\mathrm{den}\, h(\theta) \times h(\theta))}\right|.$$

Now $|N(\mathrm{den}\, h(\theta) \times h(\theta))| \ge 1$ since $(\mathrm{den}\, h(\theta) \times h(\theta))$ is a nonzero integer of $\mathbb{K}$. Consequently

$$\left|\sigma_i\left(\frac{1}{h(\theta)}\right)\right| \le (\mathrm{den}\, h(\theta))^k \times \overline{|h(\theta)|}^{k-1} \quad (1 \le i \le k),$$

which proves Lemma 3.2 by using again Lemma 3.1. $\square$

Now we prove Theorem 1.2. For every $z \in \mathbb{C}$ satisfying $|z| > 1/C_f$ and every $n \ge 0$, put

$$q(p^n(z)) = \sum_{i=0}^{ld^n} \alpha_{n,i} z^i, \quad \alpha_{n,ld^n} = b^{(d^n-1)/(d-1)} \neq 0.$$

Then

$$\frac{a^n}{q(p^n(1/z))} = \frac{a^n z^{ld^n}}{\sum_{i=0}^{ld^n} \alpha_{n,i} z^{ld^n - i}}, \tag{33}$$

so that the function

$$F(z) = f\left(\frac{1}{z}\right) = \sum_{n=0}^{\infty} \frac{a^n}{q(p^n(1/z))} \tag{34}$$

is analytic in $\mathcal{E}_f = \{z \in \mathbb{C} \mid |z| < 1/C_f\}$.

If f is algebraic over $\mathbb{C}(z)$, we have the exceptional case by Theorem 1.1. From now on let f be not algebraic over $\mathbb{C}(z)$, and the coefficients of $p(z)$ and $q(z)$ be algebraic numbers, as well as x, a, and $f(x)$. We may assume without loss of generality that $x \in D_f$, since otherwise we can choose n_0 such

that $p^n(x) \in D_f$ for all $n \geq n_0$ and consider the value $f(x')$ with $x' = p^{n_0}(x)$. To prove the theorem, we assume that the value $f(x)$ is algebraic and deduce a contradiction.

Let $\mathbb{K} \subset \mathbb{C}$ be the number field generated by all these numbers, let $\mathbb{A}$ be the ring of integers of $\mathbb{K}$, and let $k = \deg \mathbb{K}$. It is clear from (33) and (34) that the power series expansions of $F(z)$ and all its powers, namely

$$(F(z))^j = \sum_{n=0}^{\infty} \gamma_{j,n} z^n, \tag{35}$$

satisfy $\gamma_{j,n} \in \mathbb{K}$ for all nonnegative integers j and n. Now let r be a fixed positive integer. We claim that there exist polynomials $P_0, P_1, \ldots, P_r \in \mathbb{A}[z]$ of degrees at most r, not all zero, such that

$$P_0(z) + P_1(z)F(z) + P_2(z)(F(z))^2 + \cdots + P_r(z)(F(z))^r = z^{r^2+\sigma} L_r(z), \tag{36}$$

where $\sigma = \sigma(r) \geq 0$, $L_r(z) \in \mathbb{K}[\![z]\!]$ with $L_r(0) \neq 0$. Indeed, the left-hand side is not identically 0 since F is not algebraic. To realize (36) we have to solve a system of r^2 homogeneous equations (the coefficients of the successive powers z^i of the left-hand side must be equal to 0 for i from 0 to $r^2 - 1$) with $(r+1)^2$ unknowns (the coefficients of the P_i's). Since $(F(z))^h \in \mathbb{K}[\![z]\!]$ for every nonnegative integer h, we know from an elementary result of linear algebra that the system has a nontrivial solution in $\mathbb{K}^{(r+1)^2}$, and hence in $\mathbb{A}^{(r+1)^2}$ if we multiply by a common denominator, which proves our claim.

Replacing z by $1/p^n(x)$ yields

$$\theta_{r,n} = \sum_{j=0}^{r} P_j\left(\frac{1}{p^n(x)}\right)(f(p^n(x)))^j = \left(\frac{1}{p^n(x)}\right)^{r^2+\sigma} L_r\left(\frac{1}{p^n(x)}\right). \tag{37}$$

Under our hypotheses, the left-hand side of (37), which we call $\theta_{r,n}$, belongs to $\mathbb{K}$. As usual, we will obtain a contradiction by letting n tend to infinity for a suitable value of r and applying the size inequality to $\theta_{r,n}$. In what follows, we denote by $C_1, C_2, \ldots$ real numbers greater than 1 which do not depend on n or r (they may depend on x, $p(x)$ or $f(x)$).

Lemma 3.3. *There exists C_1 such that*

$$\max\left(\operatorname{den}\left(\frac{1}{q(p^n(x))}\right), \overline{\left|\frac{1}{q(p^n(x))}\right|}\right) \leq C_1^{d^n}. \tag{38}$$

Proof. An easy induction using Lemma 3.1(i) shows that, for every $n \geq 1$,

$$\operatorname{den}(p^n(x)) \leq \mu(p)^{(d^n-1)/(d-1)} (\operatorname{den} x)^{d^n} \leq C_2^{d^n}. \tag{39}$$

Furthermore, we have by Lemma 3.1(ii)

$$\overline{|p^n(x)|} \leq \mu(p)^{(d^n-1)/(d-1)} (\max(1, \overline{|x|}))^{d^n} \leq C_3^{d^n}. \tag{40}$$

For $n \geq 2$, we see by Lemma 3.2 that

$$\max\left(\operatorname{den}\left(\frac{1}{p^n(x)}\right), \overline{\left|\frac{1}{p^n(x)}\right|}\right) \leq \nu(p)\big(\operatorname{den} p^{n-1}(x) \times \max(1, \overline{|p^{n-1}(x)|})\big)^{kd}.$$

Therefore by (39) and (40)

$$\max\left(\operatorname{den}\left(\frac{1}{p^n(x)}\right), \overline{\left|\frac{1}{p^n(x)}\right|}\right) \le C_4^{d^n}. \tag{41}$$

By Lemma 3.2, this implies (38). □

Lemma 3.4. *There exist C_5, C_6, and C_7 such that*

$$\operatorname{den}(\theta_{r,n}) \le C_5^{rd^n}, \tag{42}$$

$$\overline{|\theta_{r,n}|} \le (r+1)^2 \chi C_6^{rd^n}, \tag{43}$$

$$|\theta_{r,n}| \le 2L_r(0) C_7^{-r^2 d^n}, \tag{44}$$

where χ is the greatest house of all the coefficients of all the polynomials P_i, which depends on r.

Proof. First we prove the inequality (42). By using (5), we have

$$\begin{aligned}\operatorname{den}(f(p^n(x))) &= \operatorname{den}\left(\frac{1}{a^n}\left(f(x) - \sum_{j=0}^{n-1} \frac{a^j}{q(p^j(x))}\right)\right) \\ &\le \left(\operatorname{den}\left(\frac{1}{a}\right)\right)^n \times \operatorname{den}(f(x)) \times (\operatorname{den} a)^{n-1} \times \prod_{j=0}^{n-1} \operatorname{den}\left(\frac{1}{q(p^j(x))}\right).\end{aligned}$$

By using (41), we obtain

$$\operatorname{den}(f(p^n(x))) \le C_8^n \times \prod_{j=0}^{n-1} C_1^{d^j} \le C_9^{d^n}. \tag{45}$$

The polynomials P_i defined in (36) have integer coefficients and their degrees are at most r. Hence for every $i = 0, 1, \ldots, r$, we have by (41)

$$\operatorname{LCM}\left(\operatorname{den}\left(P_i\left(\frac{1}{p^n(x)}\right)\right)\right) \le \left(\operatorname{den}\left(\frac{1}{p^n(x)}\right)\right)^r \le C_4^{rd^n}. \tag{46}$$

Now we can give an upper bound for the denominator of $(\theta_{r,n})$:

$$\operatorname{den}(\theta_{r,n}) = \operatorname{den}\left(\sum_{j=0}^{r} P_j\left(\frac{1}{p^n(x)}\right)(f(p^n(x)))^j\right) \le C_4^{rd^n} \times C_9^{rd^n} \le C_5^{rd^n}.$$

Next, we prove the inequality (43). For every $i = 0, 1, \ldots, r$, we have by (41)

$$\overline{\left|P_i\left(\frac{1}{p^n(x)}\right)\right|} \le \chi \sum_{i=0}^{r} \overline{\left|\frac{1}{p^n(x)}\right|}^i \le (r+1)\chi C_4^{rd^n}. \tag{47}$$

For every $n \ge 0$, we have by (5) and (38) above

$$\overline{|f(p^n(x))|} \le \overline{\left|\frac{1}{a^n}\right|}\left(\overline{|f(x)|} + \left(\sum_{j=0}^{n-1} \overline{|a|}^j\right) C_1^{d^n}\right) \le C_{10}^{d^n}. \tag{48}$$

By using (47) and (48), we can give an upper bound for the house of $\theta_{r,n}$:

$$\overline{|\theta_{r,n}|} \le \sum_{i=0}^{r} \overline{\left|P_i\left(\frac{1}{p^n(x)}\right)\right|} \times \overline{|[f(p^n(x))]^i|}$$
$$\le (r+1)\chi \sum_{i=0}^{r} C_4^{rd^n} \times C_{10}^{rd^n} \le (r+1)^2 \chi C_6^{rd^n}.$$

Finally, we show the inequality (44). By (37), we have

$$|\theta_{r,n}| = \left(\frac{1}{|p^n(x)|}\right)^{r^2+\sigma} \left|L_r\left(\frac{1}{p^n(x)}\right)\right|. \tag{49}$$

Since $|p^n(x)| \ge C_7^{d^n}$, we see that

$$\lim_{n\to\infty} \left|L_m\left(\frac{1}{p^n(x)}\right)\right| = |L_r(0)| \ne 0, \tag{50}$$

which proves that $\theta_{r,n} \ne 0$ for every large n. Moreover, by (49) we have (44). □

We come now to the conclusion. Define $\delta = \deg(\theta_{r,n})$. As $\theta_{r,n} \ne 0$ for every large n, it satisfies the size inequality:

$$|\theta_{r,n}| \ge (\mathrm{den}(\theta_{r,n}))^{-\delta} \times \overline{|\theta_{r,n}|}^{-\delta+1}. \tag{51}$$

Using (42), (43) and (44) yields

$$2(\gamma(r+1)^2)^{\delta-1} L_r(0) \ge \left(\frac{C_7^r}{C_5^\delta \times C_6^\delta}\right)^{rd^n}. \tag{52}$$

If we choose r such that $C_7^r > C_5^\delta \times C_6^\delta$ and fix it, we obtain a contradiction when n tends to infinity, which proves Theorem 1.2.

References

[Becker 1992] P.-G. Becker, "Algebraic independence of the values of certain series by Mahler's method", *Monatsh. Math.* **114**:3-4 (1992), 183–198. MR Zbl

[Cahen 1891] E. Cahen, "Note sur un développement des quantités numériques, qui présente quelque analogie avec celui en fractions continues", *Nouvelles Annales de Mathématiques* (3) **10** (1891), 508–514. JFM

[Davison and Shallit 1991] J. L. Davison and J. O. Shallit, "Continued fractions for some alternating series", *Monatsh. Math.* **111**:2 (1991), 119–126. MR Zbl

[Duverney 2001] D. Duverney, "Transcendence of a fast converging series of rational numbers", *Math. Proc. Cambridge Philos. Soc.* **130**:2 (2001), 193–207. MR Zbl

[Duverney et al. 2002] D. Duverney, T. Kanoko, and T. Tanaka, "Transcendence of certain reciprocal sums of linear recurrences", *Monatsh. Math.* **137**:2 (2002), 115–128. MR Zbl

[Kanoko et al. 2009] T. Kanoko, T. Kurosawa, and I. Shiokawa, "Transcendence of reciprocal sums of binary recurrences", *Monatsh. Math.* **157**:4 (2009), 323–334. MR Zbl

[Philippon 1986] P. Philippon, "Critères pour l'indépendance algébrique", *Inst. Hautes Études Sci. Publ. Math.* **64** (1986), 5–52. MR Zbl

Received 10 Jan 2018.

DANIEL DUVERNEY:

daniel.duverney@numericable.fr
Baggio Engineering School, Lille, France

TAKESHI KUROSAWA:

tkuro@rs.tus.ac.jp
Department of Applied Mathematics, Tokyo University of Science, Tokyo, Japan

IEKATA SHIOKAWA:

shiokawa@beige.ocn.ne.jp
Department of Mathematics, Keio University, Yokohama, Japan

msp

Moscow Journal of Combinatorics and Number Theory
Vol. 8, No. 1, 2019
dx.doi.org/10.2140/moscow.2019.8.71

Algebraic results for the values $\vartheta_3(m\tau)$ and $\vartheta_3(n\tau)$ of the Jacobi theta-constant

Carsten Elsner, Florian Luca and Yohei Tachiya

Let $\vartheta_3(\tau) = 1 + 2\sum_{\nu=1}^{\infty} e^{\pi i \nu^2 \tau}$ denote the classical Jacobi theta-constant. We prove that the two values $\vartheta_3(m\tau)$ and $\vartheta_3(n\tau)$ are algebraically independent over $\mathbb{Q}$ for any τ in the upper half-plane such that $q = e^{\pi i \tau}$ is an algebraic number, where $m, n \geq 2$ are distinct integers.

1. Introduction and statement of the results

Throughout this paper, let τ be a complex variable in the upper half-plane $\mathbb{H} := \{\tau \in \mathbb{C} \mid \Im(\tau) > 0\}$. The three classical theta functions

$$\vartheta_2(\tau) = 2\sum_{\nu=0}^{\infty} q^{(\nu+1/2)^2}, \quad \vartheta_3(\tau) = 1 + 2\sum_{\nu=1}^{\infty} q^{\nu^2}, \quad \vartheta_4(\tau) = 1 + 2\sum_{\nu=1}^{\infty} (-1)^{\nu} q^{\nu^2}$$

are known as theta-constants or Thetanullwerte, where $q := e^{\pi i \tau}$. These theta-constants are holomorphic in $\mathbb{H}$ and never vanish for any $\tau \in \mathbb{H}$. In particular, the function $\vartheta_3(\tau)$ is called a Jacobi theta-constant or Thetanullwert of the Jacobi theta function $\vartheta(z \mid \tau) = \sum_{\nu=-\infty}^{\infty} e^{\pi i \nu^2 \tau + 2\pi i \nu z}$. For an extensive discussion of the Jacobi theta function and theta-constants we refer the reader to [Stein and Shakarchi 2003, Chapter 10]. Y. V. Nesterenko [2006] has improved upon a result from [Grinspan 2001] and obtained some identities for the theta-constants.

Theorem A [Nesterenko 2006, Theorem 1]. *For any odd integer $n \geq 3$ there exists an integer polynomial $P_n(X, Y)$ with $\deg_X P_n(X, Y) = \psi(n)$ such that*

$$P_n\left(n^2 \frac{\vartheta_3^4(n\tau)}{\vartheta_3^4(\tau)}, 16\frac{\vartheta_2^4(\tau)}{\vartheta_3^4(\tau)}\right) = 0$$

holds for any $\tau \in \mathbb{H}$, where

$$\psi(n) := n \prod_{p \mid n} \left(1 + \frac{1}{p}\right).$$

For example, the first polynomials P_3 and P_5 are given in [Nesterenko 2006] by

$$P_3 = 9 - (28 - 16Y + Y^2)X + 30X^2 - 12X^3 + X^4,$$
$$P_5 = 25 - (126 - 832Y + 308Y^2 - 32Y^3 + Y^4)X + (255 + 1920Y - 120Y^2)X^2 + (-260 + 320Y - 20Y^2)X^3 + 135X^4 - 30X^5 + X^6$$

MSC2010: primary 11J85; secondary 11J91, 11F27.
Keywords: algebraic independence, Jacobi theta-constants, modular functions.

and the polynomials P_7, P_9, and P_{11} are listed in the appendix of [Elsner 2015]. Recently one of us (Elsner) constructed similar integer polynomials in two variables X and Y, which vanish identically at certain rational functions of theta-constants including the function $\vartheta_3(n\tau)$ for $n = 2^m$. He applied this result and Theorem A to settle the algebraic independence problem of the two values $\vartheta_3(\tau)$ and $\vartheta_3(n\tau)$ for integers $n \geq 2$, and obtained the following Theorem B.

Theorem B [Elsner 2015, Theorem 1.1]. *Let $\tau \in \mathbb{H}$ such that $e^{\pi i\tau}$ is an algebraic number. Then the two values $\vartheta_3(\tau)$ and $\vartheta_3(2^m\tau)$ are algebraically independent over $\mathbb{Q}$ for each integer $m \geq 1$. Furthermore, the same holds for the two values $\vartheta_3(\tau)$ and $\vartheta_3(n\tau)$ if $n = 3, 5, 6, 7, 9, 10, 11, 12$.*

The proof of Theorem B is based on an algebraic independence criterion, see [Elsner et al. 2011, Lemma 3.1], which requires a nonvanishing of a Jacobian determinant. In particular, to prove the latter assertion in Theorem B, he needed the explicit forms of the polynomials P_3, P_5, P_7, P_9 and P_{11} stated above. In [Elsner and Tachiya 2017], two of us obtained the following Theorem C by studying the specific properties of the polynomials P_n.

Theorem C [Elsner and Tachiya 2017, Theorem 1.2]. *Let $n \geq 2$ be an integer and $j \in \{2, 3, 4\}$. Then for any $\tau \in \mathbb{H}$ at least three of the numbers $e^{\pi i\tau}$, $\vartheta_3(\tau)$, $\vartheta_3(n\tau)$, and $D\vartheta_j(\tau)$ are algebraically independent over $\mathbb{Q}$, where $D := (\pi i)^{-1}d/d\tau$ denotes a differential operator.*

An application of Theorem C gives an improvement of Theorem B as follows:

Theorem D. *Let $\tau \in \mathbb{H}$ be such that $e^{\pi i\tau}$ is an algebraic number. Then the two numbers $\vartheta_3(\tau)$ and $\vartheta_3(n\tau)$ are algebraically independent over $\mathbb{Q}$ for each integer $n \geq 2$.*

On the other hand, the algebraic dependence result is also obtained in [Elsner and Tachiya 2017] through the properties of the polynomials P_n.

Theorem E [Elsner and Tachiya 2017, Theorem 1.4]. *Let $\ell, m, n \geq 1$ be integers and $\tau \in \mathbb{H}$ be any complex number. Then the three values $\vartheta_3(\ell\tau)$, $\vartheta_3(m\tau)$, and $\vartheta_3(n\tau)$ are algebraically dependent over $\mathbb{Q}$.*

In this paper, we fill the gap between Theorems D and E. Our main result is the following.

Theorem 1. *Let $m, n \geq 1$ be distinct integers and $\tau \in \mathbb{H}$. Then at least two of the numbers $e^{\pi i\tau}$, $\vartheta_3(m\tau)$, and $\vartheta_3(n\tau)$ are algebraically independent over $\mathbb{Q}$. In particular, the two numbers $\vartheta_3(m\tau)$ and $\vartheta_3(n\tau)$ are algebraically independent over $\mathbb{Q}$ for any $\tau \in \mathbb{H}$ such that $e^{\pi i\tau}$ is an algebraic number.*

Of course the two numbers $\vartheta_3(m\tau)$ and $\vartheta_3(n\tau)$ can be algebraically dependent over $\mathbb{Q}$ without an algebraic condition on $e^{\pi i\tau}$. Indeed, for $\tau = i \in \mathbb{H}$ the two numbers $\vartheta_3(i)$ and $\vartheta_3(2i)$ are algebraically dependent over $\mathbb{Q}$, since the nontrivial relation

$$4\vartheta_3^2(2i) - (\sqrt{2}+2)\vartheta_3^2(i) = 0 \tag{1}$$

exists; see [Berndt 1998, p. 325]. Note that the number $e^{\pi} = i^{-2i}$ was shown to be transcendental for the first time by A. O. Gelfond (1929) and, a few years later, this property of e^{π} was corroborated by the Gelfond–Schneider theorem (1934). Conversely, the above identity (1) and Theorem 1 imply the transcendence of e^{π} as well.

2. Some properties of $P_n(X, Y)$

We now discuss some properties of $P_n(X, Y)$ given in Theorem A. We start with a short description of the construction of $P_n(X, Y)$; for details, see [Nesterenko 2006]. Let $\Gamma(2)$ be the principal congruence subgroup of level 2 in $\mathrm{SL}(2, \mathbb{Z})$; that is,

$$\Gamma(2) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbb{Z}) \;\middle|\; \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod 2 \right\}.$$

Then for each odd integer $n \geq 3$ the set of matrices

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod 2, \quad (a, b, c, d) = 1, \;\; ad - bc = n,$$

is a union of $\psi(n)$ equivalence classes with respect to the left–multiplication on the elements of $\Gamma(2)$, and the class representatives are given by

$$\alpha_\nu := \begin{pmatrix} u & 2v \\ 0 & w \end{pmatrix}, \quad (u, v, w) = 1, \;\; uw = n, \;\; 0 \leq v < w. \tag{2}$$

For these $\psi(n)$ matrices $\alpha_1, \ldots, \alpha_{\psi(n)}$ in (2), we define the polynomial

$$\prod_{\nu=1}^{\psi(n)} (X - x_\nu(\tau)) =: X^{\psi(n)} + a_1(\tau) X^{\psi(n)-1} + \cdots + a_{\psi(n)-1}(\tau) X + a_{\psi(n)}(\tau),$$

where

$$x_\nu(\tau) := u^2 \frac{\vartheta_3^4((u\tau + 2v)/w)}{\vartheta_3^4(\tau)} \quad \text{with} \quad \begin{pmatrix} u & 2v \\ 0 & w \end{pmatrix} = \alpha_\nu, \;\; \nu = 1, \ldots, \psi(n). \tag{3}$$

Then, using the modular method as well as Galois considerations, one finds that there exist polynomials $R_j(Y) \in \mathbb{Z}[Y]$, $j = 1, \ldots, \psi(n)$, such that

$$a_j(\tau) = R_j(16\lambda(\tau)), \quad \lambda(\tau) := \frac{\vartheta_2^4(\tau)}{\vartheta_3^4(\tau)}. \tag{4}$$

Thus, the integer polynomial

$$P_n(X, Y) := X^{\psi(n)} + R_1(Y) X^{\psi(n)-1} + \cdots + R_{\psi(n)-1}(Y) X + R_{\psi(n)}(Y) \tag{5}$$

vanishes identically at $X = n^2 \vartheta_3^4(n\tau)/\vartheta_3^4(\tau)$ and $Y = 16\lambda(\tau)$.

Lemma 2. *For each odd integer $n \geq 3$, the polynomial $P_n(X, 16\lambda(\tau))$ is irreducible over the field $\mathbb{C}(\lambda(\tau))$.*

Proof. The group $\Gamma(2)$ fixes the function $\lambda(\tau) = \vartheta_2^4(\tau)/\vartheta_3^4(\tau)$, since the functions $\vartheta_3^4(\tau)$ and $\vartheta_4^4(\tau)$ are modular forms of weight 2 with respect of the subgroup $\Gamma(2)$. Moreover, we have the transformation formula

$$x_\nu\left(\frac{a\tau + b}{c\tau + d}\right) = x_\mu(\tau) \tag{6}$$

for a proper matrix $\beta := \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma(2)$ and subscripts ν, μ such that a proper matrix $\gamma \in \Gamma(2)$ satisfies $\alpha_\nu \beta = \gamma \alpha_\mu$; see formulae (6) and (7) in [Nesterenko 2006]. This may be regarded as an equivalence relation over

the matrices $\alpha_1, \alpha_2, \ldots, \alpha_{\psi(n)}$ from (2). One can show that any two matrices α_ν and α_μ, $1 \le \nu, \mu \le \psi(n)$, are equivalent. Together with (6) it turns out that the group $\Gamma(2)$ permutes the $\psi(n)$ distinct functions $x_1(\tau), \ldots, x_{\psi(n)}(\tau)$ transitively. This implies that $P_n(X, 16\lambda(\tau))$ is a minimal polynomial of $x_1(\tau)$ over the field $\mathbb{C}(\lambda(\tau))$. □

Remark 3. There is no complex number α such that $P_n(\alpha, Y)$ is identically zero. If such an α existed, the polynomial $P_n(X, Y)$ would be divisible by $(X - \alpha)$, which is impossible by Lemma 2. This fact can also be checked directly from the definition of $x_\nu(\tau)$; see [Elsner and Tachiya 2017, Lemma 2.1]. In particular, $P_n(X, Y)$ has positive degree in Y.

Lemma 4. *We have*

$$P_n(X, 0) = \prod_{u \mid n,\ u \ge 1} (X - u^2)^{w(u, n/u)},$$

where

$$w(a, b) := \sum_{\substack{(a,b,k)=1 \\ 0 \le k < b}} 1.$$

Proof. This follows immediately from the relation

$$P_n(X, 16\lambda(\tau)) = \prod_{\nu=1}^{\psi(n)} (X - x_\nu(\tau))$$

as $\tau \to i\infty$, since we have $\lambda(\tau) \to 0$ and $x_\nu(\tau) \to u^2$ for each $\nu = 1, \ldots, \psi(n)$ in (3), respectively. □

Example 5. For the polynomial P_3 given in Section 1, we have

$$P_3(X, 0) = 9 - 28X + 30X^2 - 12X^3 + X^4 = (X - 1)^3(X - 3^2).$$

Here, $\psi(3) = 4$ and the four triples (u, v, w) in (2) are given by

$$(3, 0, 1), \quad (1, 0, 3), \quad (1, 1, 3), \quad (1, 2, 3).$$

More generally, $P_p(X, 0) = (X - 1)^p(X - p^2)$ for any odd prime $p \ge 3$.

3. Lemmas

Let $\tau \in \mathbb{H}$. We prove in Lemmas 7 and 8 below that the number $\vartheta_3(\tau)$ is algebraic over the field $\mathbb{Q}(\vartheta_3(u\tau), \vartheta_3(v\tau))$ for certain positive integers u and v. To see this, we need the following Lemma 6. Note that $P_n(0, Y)$ is a *nonzero integer* for the polynomial $P_n(X, Y)$ in Theorem A; see [Elsner and Tachiya 2017, Lemma 2.3].

Lemma 6 [Elsner and Tachiya 2017, Lemma 2.5]. *Let $n = 2^\alpha m$ be an integer with $\alpha \ge 1$ and odd integer $m \ge 3$. Then there exists a polynomial $Q_n(X, Y) \in \mathbb{Z}[X, Y]$ such that*

$$Q_n\left(\frac{\vartheta_3^4(n\tau)}{\vartheta_3^4(\tau)}, \frac{\vartheta_2^4(\tau)}{\vartheta_3^4(\tau)}\right) = 0$$

for any $\tau \in \mathbb{H}$. Furthermore, the polynomial $Q_n(X, Y)$ is of the form

$$Q_n(X, Y) = c^{2^\alpha} Y^{2^\alpha \psi(m)} + \sum_{j=0}^{2^\alpha \psi(m)-1} R_{n,j}(X) Y^j, \tag{7}$$

with

$$Q_n(0, Y) = c^{2^\alpha} Y^{2^\alpha \psi(m)},$$

where c is equal to the nonzero integer $P_m(0, Y)$.

First we consider the case where u and v have different parity.

Lemma 7. *Let $u \geq 1$ be an odd integer and $v \geq 2$ be an even integer which is not a power of 2. Then for any $\tau \in \mathbb{H}$ the number $\vartheta_3(\tau)$ is algebraic over the field $\mathbb{Q}(\vartheta_3(u\tau), \vartheta_3(v\tau))$.*

Proof. The assertion is clear if $u = 1$. Let $u \geq 3$ be an odd integer and $P_u(X, Y)$ be as in Theorem A. Then

$$P_u\left(u^2 \frac{\vartheta_3^4(u\tau)}{\vartheta_3^4(\tau)}, 16 \frac{\vartheta_2^4(\tau)}{\vartheta_3^4(\tau)}\right) = 0 \tag{8}$$

for any $\tau \in \mathbb{H}$. Noting that $P_u(X, Y)$ has positive degree in Y and $P_u(0, Y)$ is a nonzero integer, we have the form

$$P_u(X, Y) = \sum_{j=0}^{d_u} S_{u,j}(X) Y^j, \quad S_{u,d_u}(X) \not\equiv 0,$$

with

$$c_u := S_{u,0}(0) = P_u(0, 0) \in \mathbb{Z} \setminus \{0\} \quad \text{and} \quad S_{u,j}(0) = 0 \quad (1 \leq j \leq d_u). \tag{9}$$

On the other hand, since v is not a power of 2, Lemma 6 shows that there exists a nonzero polynomial $Q_v(X, Y) \in \mathbb{Z}[X, Y]$ such that

$$Q_v\left(\frac{\vartheta_3^4(v\tau)}{\vartheta_3^4(\tau)}, \frac{\vartheta_2^4(\tau)}{\vartheta_3^4(\tau)}\right) = 0 \tag{10}$$

for any $\tau \in \mathbb{H}$, where $Q_v(X, Y)$ is of the form (7) with

$$Q_v(0, Y) := c_v Y^{d_v}, \quad c_v \in \mathbb{Z} \setminus \{0\}. \tag{11}$$

Let $\tau \in \mathbb{H}$ be a fixed complex number. Then, by (8) and (10), the polynomials $P_u(u^2\vartheta_3^4(u\tau)/\vartheta_3^4(\tau), 16Y)$ and $Q_v(\vartheta_3^4(v\tau)/\vartheta_3^4(\tau), Y)$ have the same common root $Y_0 = \vartheta_2^4(\tau)/\vartheta_3^4(\tau)$. Hence, the resultant

$$R_1(X, Z) := \mathrm{Res}_Y(P_u(X, 16Y), Q_v(Z, Y))$$

is given by the determinant D_Y of the square $(d_u + d_v)$ Sylvester matrix depending on the coefficients of $P_u(X, 16Y)$ and $Q_v(Z, Y)$ with respect to Y. Then, $R_1(X, Z)$ (and thus D_Y) vanishes at $X := u^2\vartheta_3^4(u\tau)/\vartheta_3^4(\tau)$ and $Z := \vartheta_3^4(v\tau)/\vartheta_3^4(\tau)$, so that the polynomial

$$R_2(W) := R_1(u^2\vartheta_3^4(u\tau)W, \vartheta_3^4(v\tau)W)$$

has a root $W_0 = \vartheta_3^{-4}(\tau)$ over the field $K := \mathbb{Q}(\vartheta_3(u\tau), \vartheta_3(v\tau))$. Note that $R_2(W)$ is not identically zero, since by (9) and (11) the determinant D_Y takes the form

$$R_2(0) = R_1(0,0) = \det \begin{pmatrix} & & & c_u & 0 & 0 \\ & & & & \ddots & 0 \\ & & & & & c_u \\ c_v & & & & & \\ 0 & \ddots & & & & \\ 0 & 0 & c_v & & & \end{pmatrix} = \pm c_u^{d_v} c_v^{d_u} \neq 0.$$

Therefore the number $\vartheta_3(\tau)$ is algebraic over K and the proof of Lemma 7 is completed. □

Next we treat the case where both u and v are odd.

Lemma 8. *Let $u, v \geq 1$ be distinct odd integers. Then for any $\tau \in \mathbb{H}$ the number $\vartheta_3(\tau)$ is algebraic over the field $\mathbb{Q}(\vartheta_3(u\tau), \vartheta_3(v\tau))$.*

Proof. We may assume $u, v \geq 3$. Similarly to the proof of Lemma 7, we consider the resultant

$$R_1(X, Z) := \mathrm{Res}_Y(P_u(X, Y), P_v(Z, Y)), \tag{12}$$

and the polynomial

$$R_2(W) := R_1(u^2\vartheta_3^4(u\tau)W, v^2\vartheta_3^4(v\tau)W), \tag{13}$$

which has a root $W_0 = \vartheta_3^{-4}(\tau)$. Suppose to the contrary that the above polynomial $R_2(W)$ is identically zero for some $\tau_0 \in \mathbb{H}$. Then, putting $\alpha := u^2\vartheta_3^4(u\tau_0)$ and $\beta := v^2\vartheta_3^4(v\tau_0)$, we have by (12) and (13)

$$\mathrm{Res}_Y(P_u(\alpha W, Y), P_v(\beta W, Y)) = R_1(\alpha W, \beta W) = R_2(W) \equiv 0,$$

and so there exists a common factor $H(W, Y) \in \mathbb{C}[W, Y]$ with positive degree in Y of the two polynomials $P_u(\alpha W, Y)$ and $P_v(\beta W, Y)$. Let

$$P_u(\alpha W, Y) = H(W, Y)\, G(W, Y).$$

Substituting the function $\lambda(\tau)$ defined by (4) into Y in the above, we have

$$P_u(\alpha W, 16\lambda(\tau)) = H(W, 16\lambda(\tau))\, G(W, 16\lambda(\tau)). \tag{14}$$

In what follows, we denote by $\deg H(W, Y)$, $\deg G(W, Y)$, and $\deg P_u(\alpha W, Y)$ the total degrees of the polynomials $H(W, Y)$, $G(W, Y)$, and $P_u(\alpha W, Y)$ with respect to W and Y, respectively. Then

$$\deg_W H(W, 16\lambda(\tau)) \leq \deg H(W, Y), \quad \deg_W G(W, 16\lambda(\tau)) \leq \deg G(W, Y),$$

so that

$$\begin{aligned} \deg_W P_u(\alpha W, 16\lambda(\tau)) &= \deg_W H(W, 16\lambda(\tau)) + \deg_W G(W, 16\lambda(\tau)) \\ &\leq \deg H(W, Y) + \deg G(W, Y) \\ &= \deg P_u(\alpha W, Y). \end{aligned}$$

On the other hand, it is clear that

$$\deg_W P_u(\alpha W, 16\lambda(\tau)) = \deg P_u(\alpha W, Y),$$

since by [Nesterenko 2006, Corollary 4] the inequalities

$$\deg_Y R_k(Y) \le k \cdot \frac{n-1}{n}, \quad 1 \le k \le \psi(n),$$

hold in (5). Thus, we get

$$\deg_W H(W, 16\lambda(\tau)) = \deg H(W, Y) \ge \deg_Y H(W, Y) \ge 1. \tag{15}$$

Hence by Lemma 2 together with (14) and (15), we obtain

$$P_u(\alpha W, 16\lambda(\tau)) = c_1 H(W, 16\lambda(\tau))$$

for some nonzero complex numbers c_1. Similarly there exists a nonzero complex number c_2 such that

$$P_v(\beta W, 16\lambda(\tau)) = c_2 H(W, 16\lambda(\tau)),$$

and hence

$$P_u(\alpha W, 16\lambda(\tau)) = c P_v(\beta W, 16\lambda(\tau)), \quad c := c_1/c_2.$$

Taking $\tau \to i\infty$ in the above equality, we have by Lemma 4

$$\prod_{d \mid u,\ d \ge 1} (\alpha W - d^2)^{w(d,u/d)} = c \prod_{d \mid v,\ d \ge 1} (\beta W - d^2)^{w(d,v/d)}.$$

Assume, without loss of generality, that $u > v$. Then, comparing the multiplicity of the zeros of these polynomials at $1/\alpha$, we obtain

$$u = w(1, u) \le \max_d w(d, v/d) \le v,$$

which is a contradiction. Hence, the polynomial $R_2(W)$ is not identically zero for any $\tau \in \mathbb{H}$, and the proof of Lemma 8 is completed by $R_2(\vartheta_3^{-4}(\tau)) = 0$. □

4. Proof of Theorem 1

Proof of Theorem 1. Let m and n be distinct positive integers. Define $m_1 := m/d$ and $n_1 := n/d$, where $d := \gcd(m, n)$. Without loss of generality, we may assume that m_1 is odd. In what follows, we distinguish two cases based on the parity of n_1. We first suppose that n_1 is even. Let $\tau \in \mathbb{H}$. Then, by Lemma 7 with $u := 3m_1 \ge 3$, $v := 3n_1 \ne 2^\alpha$ ($\alpha \ge 0$), and $\tau_0 := d\tau/3 \in \mathbb{H}$, the number $\vartheta_3(\tau_0)$ is algebraic over the field $\mathbb{Q}(\vartheta_3(u\tau_0), \vartheta_3(v\tau_0))$. Hence, we obtain

$$\begin{aligned}
\text{trans. deg}_{\mathbb{Q}}\, \mathbb{Q}(e^{\pi i\tau}, \vartheta_3(m\tau), \vartheta_3(n\tau)) &= \text{trans. deg}_{\mathbb{Q}}\, \mathbb{Q}(e^{\pi i\tau_0}, \vartheta_3(u\tau_0), \vartheta_3(v\tau_0)) \\
&= \text{trans. deg}_{\mathbb{Q}}\, \mathbb{Q}(e^{\pi i\tau_0}, \vartheta_3(\tau_0), \vartheta_3(u\tau_0), \vartheta_3(v\tau_0)) \\
&\ge \text{trans. deg}_{\mathbb{Q}}\, \mathbb{Q}(e^{\pi i\tau_0}, \vartheta_3(\tau_0), \vartheta_3(u\tau_0)) \\
&\ge 2,
\end{aligned}$$

where for the last inequality we used the fact that $u > 2$ and that at least two of the numbers $e^{\pi i\tau_0}$, $\vartheta_3(\tau_0)$ and $\vartheta_3(u\tau_0)$ are algebraically independent over $\mathbb{Q}$; see [Elsner and Tachiya 2017, Theorem 1.2]. In the case where n_1 is odd, we can deduce the same inequality as above by applying Lemma 8 with the same quantities u, v, τ_0 as above.

Therefore, at least two of the numbers $e^{\pi i\tau}$, $\vartheta_3(m\tau)$, and $\vartheta_3(n\tau)$ are algebraically independent over $\mathbb{Q}$, and the proof of Theorem 1 is complete. □

In the case where $m > n$ with two odd integers m, n, we obtain a stronger result based on [Elsner and Tachiya 2017, Theorem 1.2] and on Lemma 8.

Theorem 9. *Let $m > n \geq 1$ be odd integers, $j \in \{2, 3, 4\}$ and $\tau \in \mathbb{H}$. Then we have*

$$\text{trans. deg}_{\mathbb{Q}}\, \mathbb{Q}(e^{i\pi\tau}, \vartheta_3(m\tau), \vartheta_3(n\tau), D\vartheta_j(\tau)) \geq 3.$$

Proof. We apply Lemma 8 with $u = m$ and $v = n$. Therefore, we know that $\vartheta_3(\tau)$ is algebraic over the field $\mathbb{Q}(\vartheta_3(m\tau), \vartheta_3(n\tau))$. Hence we obtain with Theorem C,

$$\begin{aligned}\text{trans. deg}_{\mathbb{Q}}\, \mathbb{Q}(e^{i\pi\tau}, \vartheta_3(m\tau), \vartheta_3(n\tau), D\vartheta_j(\tau)) &= \text{trans. deg}_{\mathbb{Q}}\, \mathbb{Q}(e^{\pi i\tau}, \vartheta_3(\tau), \vartheta_3(m\tau), \vartheta_3(n\tau), D\vartheta_j(\tau)) \\ &\geq \text{trans. deg}_{\mathbb{Q}}\, \mathbb{Q}(e^{i\pi\tau}, \vartheta_3(\tau), \vartheta_3(m\tau), D\vartheta_j(\tau)) \\ &\geq 3,\end{aligned}$$

as desired. This proves the theorem. □

Acknowledgments

The authors would like to express their sincere gratitude to Professor Hirofumi Tsumura, who kindly made us realize algebraic relations for the theta values as represented by (1). The authors are also grateful to the referee for careful reading the manuscript and for giving useful comments.

This work started during a very enjoyable visit of Luca at the University of Hirosaki in January 2017 and ended during a visit of Luca to the Max Planck Institute for Mathematics in Bonn from January to July 2017. He thanks those institutions for hospitality and support. In addition, Luca was supported by grant no. CPRR160325161141 and an A-rated scientist award, both from the NRF of South Africa, and by grant no. 17-02804S of the Czech Granting Agency. Tachiya was supported by JSPS, Grant-in-Aid for Young Scientists (B), 15K17504.

References

[Berndt 1998] B. C. Berndt, *Ramanujan's notebooks, part V*, Springer, 1998. MR Zbl

[Elsner 2015] C. Elsner, "Algebraic independence results for values of theta-constants", *Funct. Approx. Comment. Math.* **52**:1 (2015), 7–27. MR Zbl

[Elsner and Tachiya 2017] C. Elsner and Y. Tachiya, "Algebraic results for certain values of the Jacobi theta-constant $\vartheta_3(\tau)$", preprint, 2017. To appear in *Math. Scand.*

[Elsner et al. 2011] C. Elsner, S. Shimomura, and I. Shiokawa, "Algebraic independence results for reciprocal sums of Fibonacci numbers", *Acta Arith.* **148**:3 (2011), 205–223. MR Zbl

[Grinspan 2001] P. Grinspan, "A measure of simultaneous approximation for quasi-modular functions", *Ramanujan J.* **5**:1 (2001), 21–45. MR Zbl

[Nesterenko 2006] Y. V. Nesterenko, "On some identities for theta-constants", pp. 151–160 in *Diophantine analysis and related fields 2006*, edited by M. Katsurada et al., Sem. Math. Sci. **35**, Keio Univ., Yokohama, 2006. MR Zbl

[Stein and Shakarchi 2003] E. M. Stein and R. Shakarchi, *Complex analysis*, Princeton Lectures in Analysis **2**, Princeton University Press, 2003. MR Zbl

Received 10 Jan 2018. Revised 18 Jun 2018.

CARSTEN ELSNER:

carsten.elsner@fhdw.de
Fachhochschule für die Wirtschaft, University of Applied Sciences, Hannover, Germany

FLORIAN LUCA:

florian.luca@wits.ac.za
School of Mathematics, University of the Witwatersrand, Johannesburg, South Africa

and

Max Planck Mathematical Institute, Bonn, Germany

and

Department of Mathematics, Faculty of Sciences, University of Ostrava, Ostrava, Czech Republic

YOHEI TACHIYA:

tachiya@hirosaki-u.ac.jp
Graduate School of Science and Technology, Hirosaki University, Hirosaki, Japan

msp

Moscow Journal of Combinatorics and Number Theory
Vol. 8, No. 1, 2019
dx.doi.org/10.2140/moscow.2019.8.81

Linear independence of 1, Li_1 and Li_2

Georges Rhin and Carlo Viola

We improve and extend the irrationality results proved by the authors (*Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **4**:3 (2005), 389–437) for dilogarithms of positive rational numbers to results of linear independence over $\mathbb{Q}$ of 1, $\mathrm{Li}_1(x)$ and $\mathrm{Li}_2(x)$ for suitable $x \in \mathbb{Q}$, both for $x > 0$ and for $x < 0$.

1. Introduction

The polylogarithm of order k, arising in Euler's work, is defined by

$$\mathrm{Li}_k(x) = \sum_{n=1}^{\infty} \frac{x^n}{n^k} \quad (|x| < 1).$$

Qualitative and quantitative irrationality results for dilogarithms $\mathrm{Li}_2\big(\frac{1}{z}\big)$ with $z \in \mathbb{Z}$, $z \in (-\infty, -5] \cup [7, +\infty)$, were proved in [Hata 1993]. Hata's results were subsequently improved and extended in [Rhin and Viola 2005], and more recently in [Viola and Zudilin 2018] and in [Marcovecchio 2016]. We showed in [Rhin and Viola 2005], henceforth abbreviated [RV05], that $\mathrm{Li}_2\big(\frac{r}{s}\big) \notin \mathbb{Q}$ for all integers r, s with $r \geq 1$ and $s \geq s_1(r) > r$, where $s_1(r)$ can be explicitly computed, and gave new irrationality measures of such $\mathrm{Li}_2\big(\frac{r}{s}\big)$. In particular we proved that $\mathrm{Li}_2\big(\frac{1}{6}\big) \notin \mathbb{Q}$, with an explicit irrationality measure, thus extending Hata's range $[7, +\infty)$ mentioned above to $[6, +\infty)$.

Viola and Zudilin [2018] proved qualitative and quantitative results of linear independence over $\mathbb{Q}$ of the four numbers

$$1, \quad \mathrm{Li}_1\Big(\frac{1}{z}\Big) = -\log\Big(1 - \frac{1}{z}\Big) = -\mathrm{Li}_1\Big(\frac{1}{1-z}\Big), \quad \mathrm{Li}_2\Big(\frac{1}{z}\Big) \quad \text{and} \quad \mathrm{Li}_2\Big(\frac{1}{1-z}\Big)$$

for all $z = \frac{s}{r}$ with integers r and s satisfying

$$r \geq 1 \quad \text{and} \quad s \geq s_2(r) > r, \tag{1-1}$$

where again $s_2(r)$ is explicit. Specifically, they proved that, for any $z = \frac{s}{r}$ satisfying (1-1) and for any $\varepsilon > 0$, there exist an effective constant $C(\varepsilon, z) > 0$, and an explicit linear independence measure $\mu(z) > 0$ over $\mathbb{Q}$ such that

$$\Big|a_0 + a_1\mathrm{Li}_1\Big(\frac{1}{z}\Big) + a_2\mathrm{Li}_2\Big(\frac{1}{z}\Big) + a_3\mathrm{Li}_2\Big(\frac{1}{1-z}\Big)\Big| > C(\varepsilon, z)\, A^{-\mu(z)-\varepsilon} \tag{1-2}$$

for all $(a_0, a_1, a_2, a_3) \in \mathbb{Z}^4 \setminus \{(0, 0, 0, 0)\}$, where $A = \max\{|a_0|, |a_1|, |a_2|, |a_3|\}$. In particular, in [Viola and Zudilin 2018] the authors proved inequalities (1-2) with $z = \frac{s}{r}$ for $r = 1$ and $s \geq 9$, for $r = 2$ and $s \geq 143$, for $r = 3$ and $s \geq 742$, for $r = 4$ and $s \geq 2355$.

MSC2010: primary 11J72; secondary 11J82, 33B30.
Keywords: polylogarithms, linear independence measures, permutation group method, saddle-point method in $\mathbb{C}^2$.

In the present paper we prove lower bounds of type (1-2) for

$$\Big|a_0 + a_1\mathrm{Li}_1\Big(\frac{1}{z}\Big) + a_2\mathrm{Li}_2\Big(\frac{1}{z}\Big)\Big|, \tag{1-3}$$

without considering $\mathrm{Li}_2(1/(1-z))$, for all $(a_0, a_1, a_2) \in \mathbb{Z}^3 \setminus \{(0,0,0)\}$ and for $z = \frac{s}{r}$ with integers r, s satisfying $r \geq 1$ and either $s \geq s_+(r) > r$, or $s \leq s_-(r) < 0$, with explicit $s_+(r)$ and $s_-(r)$. We combine some technical results employed in [RV05] and [Viola and Zudilin 2018]. In particular we use the permutation-group method, introduced by the authors in [Rhin and Viola 1996] and later used in a series of papers including [RV05] and [Viola and Zudilin 2018], and the saddle-point method in $\mathbb{C}^2$ in the form used in [Viola and Zudilin 2018]. The latter tool allows us to improve and extend the irrationality measures of $\mathrm{Li}_2\big(\frac{1}{z}\big)$ for $z = \frac{s}{r} > 1$ obtained in [RV05] to linear independence measures over $\mathbb{Q}$ of 1, $\mathrm{Li}_1\big(\frac{1}{z}\big)$ and $\mathrm{Li}_2\big(\frac{1}{z}\big)$, both for $z = \frac{s}{r} > 1$ and for $z = \frac{s}{r} < 0$. Moreover, our method yields improvements upon the linear independence measures of 1, $\mathrm{Li}_1\big(\frac{1}{z}\big)$ and $\mathrm{Li}_2\big(\frac{1}{z}\big)$ for $z < 0$ given in [Marcovecchio 2016, Table p. 231], and a new proof of the linear independence measures for $z > 1$ given therein.

We point out that, in contrast to (1-2), the lower bounds for (1-3) obtained in the present paper do not simultaneously involve values of Li_2 at positive and at negative rational numbers. Treating separately the cases $z > 1$ and $z < 0$ in (1-3), as we do in this paper, yields linear independence results for 1, $\mathrm{Li}_1\big(\frac{1}{z}\big)$ and $\mathrm{Li}_2\big(\frac{1}{z}\big)$ stronger than those obtained through the construction in [Viola and Zudilin 2018] under the constraint $a_2 = 0$ or $a_3 = 0$ in (1-2). This is not surprising, because the algebraic structure of the group $\langle \boldsymbol{\varphi}, \boldsymbol{\lambda}\rangle$ in (2-21) below generated by the permutations $\boldsymbol{\lambda}$ and $\boldsymbol{\varphi}$ in (2-18)–(2-19), used in the present paper as well as in [RV05] to get the arithmetical correction $\int_\Omega \mathrm{d}\psi(x) + \int_{\Omega'} \mathrm{d}\psi(x)$ in (4-2), is richer than that of the corresponding group $\langle \boldsymbol{\varphi}, \boldsymbol{\nu}\rangle$ required in [Viola and Zudilin 2018] to deal with the full linear form on the left-hand side of (1-2). This is a consequence of the relation $\boldsymbol{\nu}\boldsymbol{\varphi} = \boldsymbol{\varphi}\boldsymbol{\nu}$ satisfied by the permutations $\boldsymbol{\nu}$ and $\boldsymbol{\varphi}$ in [Viola and Zudilin 2018, (3.5)], which implies that $\langle \boldsymbol{\varphi}, \boldsymbol{\nu}\rangle$ is isomorphic to Klein's Vierergruppe of order 4, while the group $\langle \boldsymbol{\varphi}, \boldsymbol{\lambda}\rangle$, with $\boldsymbol{\lambda}\boldsymbol{\varphi} \neq \boldsymbol{\varphi}\boldsymbol{\lambda}$, is larger, being isomorphic to the group $\mathfrak{S}_2 \times \mathfrak{S}_3$ of order $2! \cdot 3! = 12$.

2. Simultaneous approximations to Li_1 and Li_2

Let $z \in \mathbb{R} \setminus [0,1]$, and let $h,\ j,\ k,\ l,\ m$ be nonnegative integers such that $l+m-j,\ m+h-k,\ h+j-l$ and $j+k-m$ are also nonnegative. If $z > 1$, as in [RV05, Section 2] we define the following double integrals $I_z^{(\nu)}(h,\ j,\ k,\ l,\ m)\ \ (\nu = 0, 1, 2)$:

$$I_z^{(0)}(h,\ j,\ k,\ l,\ m) = z^{-l-m}\int_0^1\int_0^1 \frac{x^j(1-x)^h y^k(1-y)^l}{(x(1-y)+yz)^{j+k-m+1}}\,\mathrm{d}x\,\mathrm{d}y, \tag{2-1}$$

$$I_z^{(1)}(h,\ j,\ k,\ l,\ m) = z^{-l-m}\int_0^1 \left(\frac{1}{2\pi i}\oint_{|y-x/(x-z)|=\varrho} \frac{x^j(1-x)^h y^k(1-y)^l}{(x(1-y)+yz)^{j+k-m+1}}\,\mathrm{d}y\right)\mathrm{d}x, \tag{2-2}$$

$$I_z^{(2)}(h,\ j,\ k,\ l,\ m) = \frac{z^{-l-m}}{2\pi i}\oint_{|x-z|=\sigma}\left(\frac{1}{2\pi i}\oint_{|y-x/(x-z)|=\varrho} \frac{x^j(1-x)^h y^k(1-y)^l}{(x(1-y)+yz)^{j+k-m+1}}\,\mathrm{d}y\right)\mathrm{d}x \tag{2-3}$$

for any $\varrho, \sigma > 0$. We also define the linear combination of (2-1) and (2-2) given by

$$I_z(h,\ j,\ k,\ l,\ m) = I_z^{(0)}(h,\ j,\ k,\ l,\ m) - (\log z)\, I_z^{(1)}(h,\ j,\ k,\ l,\ m), \tag{2-4}$$

where $\log z$ is the real value of the logarithm for $z > 1$.

Clearly the definitions (2-2) and (2-3) make sense also for $z<0$, whereas the definition (2-1) of $I_z^{(0)}(h,j,k,l,m)$ does not apply if $z<0$, since in this case the denominator $x(1-y)+yz$ vanishes for (x,y) along a segment of hyperbola inside the unit square $(0,1)\times(0,1)\subset\mathbb{R}^2$. In order to define $I_z^{(0)}(h,j,k,l,m)$ for $z<0$ we use a method introduced in [Viola and Zudilin 2018, Section 2.2]. We apply to $I_z^{(0)}(h,j,k,l,m)$ the change of variables

$$x=\xi,\quad y=\frac{\eta}{\eta-z}. \tag{2-5}$$

For $z>1$, (2-5) changes the integration path $[0,1)$ for y to $[0,-\infty)$ for η. Thus

$$I_z^{(0)}(h,j,k,l,m)=(-1)^{j+k+l+m}\,z^{-j-k}\int_0^1\xi^j(1-\xi)^h\,\mathrm{d}\xi\int_0^{-\infty}\frac{\eta^k}{(\xi-\eta)^{j+k-m+1}(\eta-z)^{l+m-j+1}}\,\mathrm{d}\eta.$$

Let ζ be any complex number such that $|\zeta|=1$, $\zeta\neq1$, and let $\arg\zeta$ denote the argument satisfying $0<\arg\zeta<2\pi$. For $\varrho>0$ let μ_ϱ be the arc $\{|\eta|=\varrho:\arg\eta \text{ from } \arg\zeta \text{ to } \pi\}$. For any ξ,z with $0<\xi\le1<z$, as $\varrho\to+\infty$ we get

$$\left|\int_{\mu_\varrho}\frac{\eta^k}{(\xi-\eta)^{j+k-m+1}(\eta-z)^{l+m-j+1}}\,\mathrm{d}\eta\right|\le\frac{\varrho^k}{(\varrho-1)^{j+k-m+1}(\varrho-z)^{l+m-j+1}}\cdot2\pi\varrho\ll\varrho^{-l-1}\to0. \tag{2-6}$$

Therefore, by Cauchy's theorem, for any $z>1$ and for any $\zeta\in\mathbb{C}$ such that $|\zeta|=1$, $\zeta\neq1$, we obtain

$$I_z^{(0)}(h,j,k,l,m)=(-1)^{j+k+l+m}\,z^{-j-k}\int_0^1\xi^j(1-\xi)^h\,\mathrm{d}\xi\int_0^{\zeta\infty}\frac{\eta^k}{(\xi-\eta)^{j+k-m+1}(\eta-z)^{l+m-j+1}}\,\mathrm{d}\eta, \tag{2-7}$$

where the integration path for η is the half-line $[0,\zeta\infty)$ from 0 to ∞ through ζ. Note that, since $z>1$ and $\zeta\neq1$, we have $(\xi-\eta)(\eta-z)\neq0$ in (2-7).

As in [Viola and Zudilin 2018, Section 2.2], it is easy to prove that the double integral on the right-hand side of (2-7) converges absolutely and uniformly when z varies in any compact region of $\mathbb{C}$ not intersecting the half-line $[0,\zeta\infty)$, and in particular the integrations in ξ and η can be interchanged. Thus

$$I_z^{(0)}(h,j,k,l,m)=(-1)^{j+k+l+m}\,z^{-j-k}\int_0^{\zeta\infty}\left(\int_0^1\frac{\xi^j(1-\xi)^h\eta^k}{(\xi-\eta)^{j+k-m+1}(\eta-z)^{l+m-j+1}}\,\mathrm{d}\xi\right)\mathrm{d}\eta, \tag{2-8}$$

and $I_z^{(0)}(h,j,k,l,m)$, viewed as a function of the complex variable z, is holomorphic in the cut plane

$$\mathbb{C}\setminus[0,\zeta\infty).$$

Therefore, if in (2-7) or (2-8) we choose $\zeta\in\mathbb{C}$ with $|\zeta|=1$, $\zeta\neq\pm1$, by analytic continuation we can move z from the half-line $z>1$ to the half-line $z<0$ along a path contained in the lower half-plane $\operatorname{Im}z<0$ if $\operatorname{Im}\zeta>0$, or in the upper half-plane $\operatorname{Im}z>0$ if $\operatorname{Im}\zeta<0$, because such a path does not cross the cut $[0,\zeta\infty)$.

Since, from (2-7),

$$\overline{I_z^{(0)}(h,j,k,l,m)}=(-1)^{j+k+l+m}\,\bar z^{-j-k}\int_0^1\xi^j(1-\xi)^h\,\mathrm{d}\xi\int_0^{\bar\zeta\infty}\frac{\eta^k}{(\xi-\eta)^{j+k-m+1}(\eta-\bar z)^{l+m-j+1}}\,\mathrm{d}\eta,$$

for any $z<0$ we get two values of $I_z^{(0)}(h,j,k,l,m)$, conjugate to each other, one defined by (2-7) or (2-8) for $\operatorname{Im}\zeta>0$, and the other for $\operatorname{Im}\zeta<0$. Accordingly, for $z<0$ we define $I_z(h,j,k,l,m)$ by (2-4),

where

$$\log z = \begin{cases} \log|z| - \pi i & \text{if } \operatorname{Im}\zeta > 0 \text{ in } I_z^{(0)}(h,\, j, k, l, m), \\ \log|z| + \pi i & \text{if } \operatorname{Im}\zeta < 0 \text{ in } I_z^{(0)}(h,\, j, k, l, m). \end{cases} \tag{2-9}$$

By applying the change of variables (2-5) to the double integrals (2-2) and (2-3) we easily get, similarly to [Viola and Zudilin 2018, (2.11) and (2.12)],

$$I_z^{(1)}(h,\, j, k, l, m) = (-1)^{j+k+l+m}\, \frac{z^{-j-k}}{2\pi i} \oint_{\Gamma_{0,1}} \left(\int_0^1 \frac{\xi^j (1-\xi)^h \eta^k}{(\xi-\eta)^{j+k-m+1} (\eta - z)^{l+m-j+1}}\, \mathrm{d}\xi \right) \mathrm{d}\eta, \tag{2-10}$$

where $\Gamma_{0,1}$ denotes any closed contour for η enclosing the interval $(0, 1)$ but not enclosing z, and, for any $\varrho_1, \varrho_2 > 0$,

$$\begin{aligned} &I_z^{(2)}(h,\, j, k, l, m) \\ &\quad = (-1)^{j+k+l+m+1} \frac{z^{-j-k}}{2\pi i} \oint_{|\eta - z| = \varrho_1} \left(\frac{1}{2\pi i} \oint_{|\xi - \eta| = \varrho_2} \frac{\xi^j (1-\xi)^h \eta^k}{(\xi-\eta)^{j+k-m+1} (\eta - z)^{l+m-j+1}}\, \mathrm{d}\xi \right) \mathrm{d}\eta. \end{aligned} \tag{2-11}$$

We showed in [RV05, (2.5) and (2.6)] that the involution $\lambda = \lambda_{x,z} : y \leftrightarrow \tilde{y}$ defined by

$$\tilde{y} = \frac{x(1-y)}{x(1-y) + yz}, \tag{2-12}$$

used as a change of variable for y in the double integrals (2-1), (2-2) and (2-3), transforms the quantities (2-1), (2-2), (2-3), and hence (2-4), according to the permutation $\boldsymbol{\lambda}$ of the exponents defined by

$$\boldsymbol{\lambda} = (j \;\; m)(k \;\; l), \tag{2-13}$$

which acts identically on h. Thus (2-2) and (2-3), and hence (2-10) and (2-11), are invariant under the action of $\boldsymbol{\lambda}$ both for $z > 1$ and for $z < 0$, and the same holds for (2-1) if $z > 1$.

It is easy to see that $I_z^{(0)}(h,\, j, k, l, m)$ is invariant under the action of the permutation $\boldsymbol{\lambda}$ also in the case $z < 0$, where (2-1) is replaced by the definition (2-7) or (2-8) for $|\zeta| = 1,\ \zeta \neq \pm 1$. Indeed, the change of variables (2-5) transforms the involution (2-12) into the involution $\eta \leftrightarrow \tilde{\eta}$ given by

$$\tilde{\eta} = z\, \frac{\xi}{\eta}. \tag{2-14}$$

If in (2-7) we use the change of variable (2-14) for η, we get

$$\begin{aligned} &I_z^{(0)}(h,\, j, k, l, m) \\ &\qquad = (-1)^{j+k+l+m}\, z^{-l-m} \int_0^1 \xi^m (1-\xi)^h\, \mathrm{d}\xi \int_0^{-\bar{\zeta}\infty} \frac{\tilde{\eta}^{\,l}}{(\xi - \tilde{\eta})^{l+m-j+1} (\tilde{\eta} - z)^{j+k-m+1}}\, \mathrm{d}\tilde{\eta}. \end{aligned} \tag{2-15}$$

Since $\operatorname{Im}(-\bar{\zeta}) = \operatorname{Im}\zeta$, by the same argument as in (2-6) we see that the right-hand side of (2-15) equals

$$(-1)^{j+k+l+m}\, z^{-l-m} \int_0^1 \xi^m (1-\xi)^h\, \mathrm{d}\xi \int_0^{\zeta\infty} \frac{\tilde{\eta}^{\,l}}{(\xi - \tilde{\eta})^{l+m-j+1} (\tilde{\eta} - z)^{j+k-m+1}}\, \mathrm{d}\tilde{\eta}.$$

By (2-7), this is $I_z^{(0)}(h, m, l, k,\, j)$. Hence for $z < 0$ we get

$$I_z^{(0)}(h,\, j, k, l, m) = I_z^{(0)}(h, m, l, k,\, j)$$

both for $\operatorname{Im}\zeta > 0$ and for $\operatorname{Im}\zeta < 0$, as claimed.

Throughout this paper, for any integer $n \geq 1$ we denote by d_n the least common multiple of $1, \dots, n$, and we set $d_0 = 1$. Also, as in [RV05, (2.9)], we define

$$\begin{aligned} H &= \max\{l+m-j,\ m+h-k,\ h+j-l,\ j+k-m\},\\ K &= \max\{l+m-j,\ \min\{m+h-k,\ h+j-l\},\ j+k-m\},\\ \alpha &= \max\{j+k,\ k+l,\ l+m\},\\ \beta &= \max\{0,\ k+l-h\},\\ \delta &= \alpha+\beta+h-k-l \end{aligned} \tag{2-16}$$

(note that the integer

$$\delta = \max\{h,\ m+h-k,\ h+j-l,\ j+k,\ k+l,\ l+m\}$$

defined in [RV05, (2.9)] equals $\alpha+\beta+h-k-l$ by virtue of Lemma 2.8 of the same paper). Clearly the integers (2-16) are nonnegative and invariant under the action of the permutation $\boldsymbol{\lambda}$ in (2-13).

We prove the following:

Theorem 2.1. *For any $z \in \mathbb{R} \setminus [0,1]$, define* (2-2), (2-3), (2-7) *with $|\zeta| = 1,\ \zeta \neq \pm 1$,* (2-9), (2-4) *and* (2-16). *Then*

$$\begin{aligned} d_H d_K z^\alpha (z-1)^\beta I_z(h,j,k,l,m) &= P(z) - Q(z)\,\mathrm{Li}_2\Big(\frac{1}{z}\Big),\\ d_H d_K z^\alpha (z-1)^\beta I_z^{(1)}(h,j,k,l,m) &= R(z) - Q(z)\,\mathrm{Li}_1\Big(\frac{1}{z}\Big),\\ d_H d_K z^\alpha (z-1)^\beta I_z^{(2)}(h,j,k,l,m) &= Q(z), \end{aligned}$$

where

$$P(z), Q(z), R(z) \in \mathbb{Z}[z], \qquad \max\{\deg P(z), \deg Q(z), \deg R(z)\} \leq \delta.$$

Proof. If $z > 1$, the theorem is [RV05, Theorem 2.1]. If $z < 0$, the theorem follows from the case $z > 1$ by analytic continuation, moving z from the half-line $z > 1$ to the half-line $z < 0$ along a path contained in the lower half-plane $\mathrm{Im}\, z < 0$ if $\mathrm{Im}\, \zeta > 0$, or in the upper half-plane $\mathrm{Im}\, z > 0$ if $\mathrm{Im}\, \zeta < 0$. □

Let

$$\mathcal{S} = \{h,\ j,\ k,\ l,\ m,\ l+m-j,\ m+h-k,\ h+j-l,\ j+k-m\}. \tag{2-17}$$

The action on the set $\mathcal{S}$ of the permutation $\boldsymbol{\lambda}$ defined in (2-13) is

$$\boldsymbol{\lambda} = (j\ \ m)(k\ \ l)(l+m-j\ \ \ j+k-m)(m+h-k\ \ \ h+j-l). \tag{2-18}$$

As in [RV05], we also consider the permutation $\boldsymbol{\varphi}$ whose action on $\mathcal{S}$ is

$$\boldsymbol{\varphi} = (h\ \ m+h-k)(j\ \ \ j+k-m)(k\ \ m). \tag{2-19}$$

We proved in [RV05, Section 3] that for any $z > 1$ the quotients

$$\frac{I_z^{(\nu)}(h,j,k,l,m)}{h!\,j!\,k!\,l!\,m!} \quad (\nu = 0,1,2) \qquad \text{and} \qquad \frac{I_z(h,j,k,l,m)}{h!\,j!\,k!\,l!\,m!} \tag{2-20}$$

are invariant under the action of $\boldsymbol{\varphi}$. Hence, by analytic continuation, the same holds in the case $z < 0$, where $I_z^{(0)}(h, j, k, l, m)$ and $\log z$ are given by (2-7) and (2-9) with $\zeta \in \mathbb{C}$ such that $|\zeta| = 1,\ \zeta \neq \pm 1$. Thus the quotients (2-20) are invariant under the action of the whole permutation group

$$\boldsymbol{\Phi} = \langle \boldsymbol{\varphi}, \boldsymbol{\lambda} \rangle \tag{2-21}$$

generated by $\boldsymbol{\lambda}$ and $\boldsymbol{\varphi}$, both for $z > 1$ and for $z < 0$. As we proved in [RV05, Section 3], $\boldsymbol{\Phi}$ is isomorphic to the product $\mathfrak{S}_2 \times \mathfrak{S}_3$ of the symmetric groups of orders 2! and 3!, whence

$$|\boldsymbol{\Phi}| = 2! \cdot 3! = 12.$$

Therefore, the transformation formulae in [RV05, p. 416] hold for $I_z^{(\nu)}(h, j, k, l, m)$ $(\nu = 0, 1, 2)$ and for $I_z(h, j, k, l, m)$, in both cases $z > 1$ and $z < 0$.

The integers α, β and δ in (2-16) are also invariant under the action of the permutation $\boldsymbol{\varphi}$ in (2-19), and hence are invariant under the action of the whole permutation group (2-21), while H and K in (2-16) are not invariant under the action of $\boldsymbol{\varphi}$. Therefore, in place of H and K, we require the integers M and N defined in [RV05, (4.1) and (4.2)], namely

$$M = \max \mathcal{S}$$

and

$$N = \max\{\max'(h,\ m+h-k,\ h+j-l),\ j,\ k,\ l,\ m,\ l+m-j,\ j+k-m\},$$

where $\max'$ denotes the second maximum in a finite sequence of real numbers. Clearly M and N are invariant under the action of the permutation group $\boldsymbol{\Phi}$ in (2-21). Also $H \leq M$ and $K \leq N$, whence Theorem 2.1 holds with M in place of H and N in place of K. Moreover, as in [RV05, p. 417], for any permutation $\boldsymbol{\chi} \in \boldsymbol{\Phi}$, Theorem 2.1 holds with h, j, k, l, m respectively replaced by $\boldsymbol{\chi}(h)$, $\boldsymbol{\chi}(j)$, $\boldsymbol{\chi}(k)$, $\boldsymbol{\chi}(l)$, $\boldsymbol{\chi}(m)$, with polynomials $P_{\boldsymbol{\chi}}(z)$, $Q_{\boldsymbol{\chi}}(z)$, $R_{\boldsymbol{\chi}}(z) \in \mathbb{Z}[z]$ depending on the left coset $\boldsymbol{\chi}\boldsymbol{\Lambda}$, where $\boldsymbol{\Lambda} = \langle \boldsymbol{\lambda} \rangle$ is the subgroup of $\boldsymbol{\Phi}$ of order 2 generated by (2-18), and with M, N, α, β, δ all independent of $\boldsymbol{\chi}$.

For h, j, k, l, m fixed and $n = 1, 2, 3, \ldots$, we replace the tuple (h, j, k, l, m) by (hn, jn, kn, ln, mn). Then Theorem 2.1 yields

$$\begin{aligned}
d_{Mn} d_{Nn} z^{\alpha n} (z-1)^{\beta n} I_z(hn, jn, kn, ln, mn) &= P_n(z) - Q_n(z)\mathrm{Li}_2\Big(\frac{1}{z}\Big),\\
d_{Mn} d_{Nn} z^{\alpha n} (z-1)^{\beta n} I_z^{(1)}(hn, jn, kn, ln, mn) &= R_n(z) - Q_n(z)\mathrm{Li}_1\Big(\frac{1}{z}\Big),\\
d_{Mn} d_{Nn} z^{\alpha n} (z-1)^{\beta n} I_z^{(2)}(hn, jn, kn, ln, mn) &= Q_n(z),
\end{aligned} \tag{2-22}$$

with polynomials $P_n(z)$, $Q_n(z)$, $R_n(z) \in \mathbb{Z}[z]$ of degrees not exceeding δn. Similarly, for any permutation $\boldsymbol{\chi} \in \boldsymbol{\Phi}$,

$$\begin{aligned}
d_{Mn} d_{Nn} z^{\alpha n} (z-1)^{\beta n} I_z(\boldsymbol{\chi}(h)n, \boldsymbol{\chi}(j)n, \boldsymbol{\chi}(k)n, \boldsymbol{\chi}(l)n, \boldsymbol{\chi}(m)n) &= P_{\boldsymbol{\chi},n}(z) - Q_{\boldsymbol{\chi},n}(z)\mathrm{Li}_2\Big(\frac{1}{z}\Big),\\
d_{Mn} d_{Nn} z^{\alpha n} (z-1)^{\beta n} I_z^{(1)}(\boldsymbol{\chi}(h)n, \boldsymbol{\chi}(j)n, \boldsymbol{\chi}(k)n, \boldsymbol{\chi}(l)n, \boldsymbol{\chi}(m)n) &= R_{\boldsymbol{\chi},n}(z) - Q_{\boldsymbol{\chi},n}(z)\mathrm{Li}_1\Big(\frac{1}{z}\Big),\\
d_{Mn} d_{Nn} z^{\alpha n} (z-1)^{\beta n} I_z^{(2)}(\boldsymbol{\chi}(h)n, \boldsymbol{\chi}(j)n, \boldsymbol{\chi}(k)n, \boldsymbol{\chi}(l)n, \boldsymbol{\chi}(m)n) &= Q_{\boldsymbol{\chi},n}(z),
\end{aligned} \tag{2-23}$$

with polynomials $P_{\boldsymbol{\chi},n}(z)$, $Q_{\boldsymbol{\chi},n}(z)$, $R_{\boldsymbol{\chi},n}(z) \in \mathbb{Z}[z]$ of degrees not exceeding δn.

By the invariance of the quotients (2-20) under the action of the group $\boldsymbol{\Phi}$, from (2-22) and (2-23) we obtain the identity

$$(\chi(h)n)!\,(\chi(j)n)!\,(\chi(k)n)!\,(\chi(l)n)!\,(\chi(m)n)!\,Q_n(z) = (hn)!\,(jn)!\,(kn)!\,(ln)!\,(mn)!\,Q_{\chi,n}(z), \quad (2\text{-}24)$$

whence

$$(\chi(h)n)!\,(\chi(j)n)!\,(\chi(k)n)!\,(\chi(l)n)!\,(\chi(m)n)!\,P_n(z) = (hn)!\,(jn)!\,(kn)!\,(ln)!\,(mn)!\,P_{\chi,n}(z), \quad (2\text{-}25)$$

$$(\chi(h)n)!\,(\chi(j)n)!\,(\chi(k)n)!\,(\chi(l)n)!\,(\chi(m)n)!\,R_n(z) = (hn)!\,(jn)!\,(kn)!\,(ln)!\,(mn)!\,R_{\chi,n}(z). \quad (2\text{-}26)$$

Let Ω and Ω' be the sets of real numbers $\omega \in [0, 1)$ defined in [RV05, p. 420], and let

$$\Delta_n = \prod_{\substack{p>\sqrt{Mn}\\ \{n/p\}\in\Omega}} p, \quad \Delta'_n = \prod_{\substack{p>\sqrt{Mn}\\ \{n/p\}\in\Omega'}} p \quad (n = 1, 2, 3, \dots),$$

where p denotes a prime number. By applying to the coefficients of the polynomials in (2-24), (2-25) and (2-26) the discussion in [RV05, p. 418–420], we see that $\Delta_n\Delta'_n$ divides all the coefficients of $P_n(z)$, $Q_n(z)$ and $R_n(z)$. Therefore

$$P^*_n(z) := (\Delta_n\Delta'_n)^{-1}P_n(z) \in \mathbb{Z}[z],$$
$$Q^*_n(z) := (\Delta_n\Delta'_n)^{-1}Q_n(z) \in \mathbb{Z}[z],$$
$$R^*_n(z) := (\Delta_n\Delta'_n)^{-1}R_n(z) \in \mathbb{Z}[z].$$

Let

$$D_n = \frac{d_{Mn}d_{Nn}}{\Delta_n\Delta'_n}.$$

Dividing the identities (2-22) by $\Delta_n\Delta'_n$ we get

$$\begin{aligned} D_n z^{\alpha n}(z-1)^{\beta n} I_z(hn,\, jn, kn, ln, mn) &= P^*_n(z) - Q^*_n(z)\mathrm{Li}_2\Big(\frac{1}{z}\Big),\\ D_n z^{\alpha n}(z-1)^{\beta n} I^{(1)}_z(hn,\, jn, kn, ln, mn) &= R^*_n(z) - Q^*_n(z)\mathrm{Li}_1\Big(\frac{1}{z}\Big),\\ D_n z^{\alpha n}(z-1)^{\beta n} I^{(2)}_z(hn,\, jn, kn, ln, mn) &= Q^*_n(z), \end{aligned} \quad (2\text{-}27)$$

with

$$P^*_n(z),\, Q^*_n(z),\, R^*_n(z) \in \mathbb{Z}[z], \quad \max\{\deg P^*_n(z), \deg Q^*_n(z), \deg R^*_n(z)\} \le \delta n. \quad (2\text{-}28)$$

By [RV05, (4.13)],

$$\lim_{n\to\infty}\frac{1}{n}\log D_n = M + N - \left(\int_\Omega \mathrm{d}\psi(x) + \int_{\Omega'}\mathrm{d}\psi(x)\right), \quad (2\text{-}29)$$

where $\psi(x) = \Gamma'(x)/\Gamma(x)$ is the logarithmic derivative of the Euler gamma-function.

3. The saddle-point method

We shall employ asymptotic formulae, as $n \to \infty$, for the linear forms involving $\mathrm{Li}_1\big(\frac{1}{z}\big)$ and $\mathrm{Li}_2\big(\frac{1}{z}\big)$ and for their common coefficients arising from (2-27). Such asymptotic formulae can be obtained by

applying to the integrals in (2-27) the saddle-point method for double complex integrals [Hata 2000, Section 1], in the form used in [Viola and Zudilin 2018].

We henceforth assume the integers belonging to the set $\mathcal{S}$ in (2-17) to be all strictly positive. Moreover, we assume

$$j \le h \quad \text{and} \quad k < m. \tag{3-1}$$

Let, as in [RV05, (5.1)],

$$f_z(x, y) = \frac{x^j (1-x)^h y^k (1-y)^l}{(x(1-y) + yz)^{j+k-m}},$$

and let

$$F_z(\xi, \eta) = \frac{\xi^j (1-\xi)^h \eta^k}{(\xi-\eta)^{j+k-m} (\eta - z)^{l+m-j}}. \tag{3-2}$$

The substitution (2-5) yields

$$(-z)^{-l-m} f_z\left(\xi, \frac{\eta}{\eta - z}\right) = (-z)^{-j-k} F_z(\xi, \eta), \tag{3-3}$$

whence

$$\begin{aligned} (-z)^{-j-k} \frac{\partial F_z}{\partial \xi} &= (-z)^{-l-m} \left.\frac{\partial f_z}{\partial x}\right|_{x=\xi,\ y=\eta/(\eta-z)}, \\ (-z)^{-j-k} \frac{\partial F_z}{\partial \eta} &= \frac{(-z)^{-l-m+1}}{(\eta-z)^2} \left.\frac{\partial f_z}{\partial y}\right|_{x=\xi,\ y=\eta/(\eta-z)}. \end{aligned} \tag{3-4}$$

Owing to (2-8), (2-10) and (2-11), the double integrals occurring in (2-27) can be written as

$$I_z^{(0)}(hn, jn, kn, ln, mn) = \pm z^{-(j+k)n} \int_0^{\zeta\infty} \left(\int_0^1 F_z(\xi, \eta)^n \frac{\mathrm{d}\xi}{\xi - \eta} \right) \frac{\mathrm{d}\eta}{\eta - z}, \tag{3-5}$$

$$I_z^{(1)}(hn, jn, kn, ln, mn) = \pm \frac{z^{-(j+k)n}}{2\pi i} \oint_{\Gamma_{0,1}} \left(\int_0^1 F_z(\xi, \eta)^n \frac{\mathrm{d}\xi}{\xi - \eta} \right) \frac{\mathrm{d}\eta}{\eta - z}, \tag{3-6}$$

$$I_z^{(2)}(hn, jn, kn, ln, mn) = \pm \frac{z^{-(j+k)n}}{2\pi i} \oint_{|\eta - z| = \varrho_1} \left(\frac{1}{2\pi i} \oint_{|\xi - \eta| = \varrho_2} F_z(\xi, \eta)^n \frac{\mathrm{d}\xi}{\xi - \eta} \right) \frac{\mathrm{d}\eta}{\eta - z}. \tag{3-7}$$

Since

$$\begin{aligned} \frac{1}{F_z} \frac{\partial F_z}{\partial \xi} &= \frac{\partial}{\partial \xi} \log F_z = \frac{j}{\xi} - \frac{h}{1-\xi} - \frac{j+k-m}{\xi - \eta}, \\ \frac{1}{F_z} \frac{\partial F_z}{\partial \eta} &= \frac{\partial}{\partial \eta} \log F_z = \frac{k}{\eta} + \frac{j+k-m}{\xi - \eta} - \frac{l+m-j}{\eta - z}, \end{aligned}$$

the saddle-points of $F_z(\xi, \eta)$, i.e., the stationary points of $F_z(\xi, \eta)$ satisfying $F_z(\xi, \eta) \neq 0$, are the solutions of the system

$$\begin{cases} \dfrac{j}{\xi} - \dfrac{h}{1-\xi} = \dfrac{j+k-m}{\xi - \eta}, \\ \dfrac{l+m-j}{\eta - z} - \dfrac{k}{\eta} = \dfrac{j+k-m}{\xi - \eta}. \end{cases} \tag{3-8}$$

As in [Viola and Zudilin 2018, (4.3)], the first equation (3-8) yields

$$\eta = H(\xi) := \xi\,\frac{(m+h-k)\xi+k-m}{(h+j)\xi-j}. \tag{3-9}$$

Subtracting the equations (3-8) and then substituting (3-9), we get the same cubic equation in ξ as in [RV05, (5.4)], namely

$$U(\xi) := \xi((m+h-k)\xi+k-m)((h+j-l)\xi+l-j) - z((h+j)\xi-j)((h+m)\xi-m) = 0. \tag{3-10}$$

Thus, denoting by (ξ_ν, η_ν) $(\nu = 0, 1, 2)$ the saddle-points of $F_z(\xi, \eta)$, we see that ξ_0, ξ_1, ξ_2 are the roots of (3-10), and

$$\eta_\nu = H(\xi_\nu) = \xi_\nu\,\frac{(m+h-k)\xi_\nu+k-m}{(h+j)\xi_\nu-j} \quad (\nu = 0, 1, 2). \tag{3-11}$$

Also, by (3-4),

$$x_\nu = \xi_\nu, \quad y_\nu = \frac{\eta_\nu}{\eta_\nu - z} \quad (\nu = 0, 1, 2), \tag{3-12}$$

where (x_ν, y_ν) are the saddle-points of $f_z(x, y)$.

We refer to the detailed discussion in [Viola and Zudilin 2018, Section 4] for the application of the saddle-point method in $\mathbb{C}^2$ to the double integrals (3-5), (3-6), (3-7). Let

$$\xi_\pm = \frac{j}{h+j} \pm \frac{\sqrt{hj(m+h-k)(j+k-m)}}{(h+j)(m+h-k)}$$

be the solutions of $\mathrm{d}H/\mathrm{d}\xi = 0$, and let

$$\eta_\pm = \frac{h(j+k-m)+j(m+h-k) \pm 2\sqrt{hj(m+h-k)(j+k-m)}}{(h+j)^2}.$$

The function (3-9) satisfies

$$\eta_+ = H(\xi_+), \quad \eta_- = H(\xi_-),$$

and maps both the upper and the lower half-circumference of diameter $[\xi_-, \xi_+]$ in the plane of the complex variable ξ onto the real interval $[\eta_-, \eta_+]$. Thus, if we denote by $\mathcal{C}$ and $\mathcal{D}$ the upper and lower half-planes of the complex variable η,

$$\mathcal{C} = \{\operatorname{Im}\eta > 0\}, \quad \mathcal{D} = \{\operatorname{Im}\eta < 0\},$$

and in the plane of the complex variable ξ we define the regions

$$\begin{aligned}
\mathcal{C}_1 &= \{\operatorname{Im}\xi > 0,\ |\xi - j/(h+j)| > R\},\\
\mathcal{D}_1 &= \{\operatorname{Im}\xi < 0,\ |\xi - j/(h+j)| > R\},\\
\mathcal{C}_2 &= \{\operatorname{Im}\xi < 0,\ |\xi - j/(h+j)| < R\},\\
\mathcal{D}_2 &= \{\operatorname{Im}\xi > 0,\ |\xi - j/(h+j)| < R\},
\end{aligned}$$

where

$$R = \frac{\sqrt{hj(m+h-k)(j+k-m)}}{(h+j)(m+h-k)},$$

we see that (3-9) is a one-to-one mapping of both $\mathcal{C}_1$ and $\mathcal{C}_2$ onto $\mathcal{C}$, and of both $\mathcal{D}_1$ and $\mathcal{D}_2$ onto $\mathcal{D}$.

In Sections 3.1 and 3.2 we shall treat the cases $z > 1$ and $z < 0$, respectively.

3.1. ***The case z > 1.*** As is shown in [RV05, (5.6)] the roots of (3-10) are real, and $\xi_0 < \xi_1 < \xi_2$ using the notation therein. By (3-1) we get

$$0 < \xi_- < \frac{m-k}{m+h-k} < \xi_0 < \frac{j}{h+j} < \xi_+ < 1 < z < \xi_2.$$

Hence, by (3-12) and [RV05, (5.15)],

$$\eta_0 < 0 < \eta_- < \eta_+ < 1 < z < \eta_2 < \xi_2.$$

Let $\xi = \Xi(\eta)$ be a local inverse of (3-9), holomorphic in an open region Δ contained in the plane of the complex variable η. We use the notation in [Viola and Zudilin 2018, Sections 4.1 and 4.2], where (ξ_*, η_*) denotes the relevant saddle-point for the double integral considered. Here we apply the saddle-point method to the double contour integral (3-7). We choose

$$\Delta = \mathcal{C} \cup \mathcal{D} \cup (\eta_+, +\infty), \quad \Xi : \Delta \to \mathcal{C}_1 \cup \mathcal{D}_1 \cup (\xi_+, +\infty), \quad (\xi_*, \eta_*) = (\xi_2, \eta_2).$$

We change the integration path $|\eta - z| = \varrho_1$ in (3-7) to a closed contour $\Gamma \subset \Delta$ of steepest descent for $|F_z(\Xi(\eta), \eta)|$ enclosing z and passing through 1 and η_2, whence

$$\max_{\eta \in \Gamma} |F_z(\Xi(\eta), \eta)| = |F_z(\xi_2, \eta_2)|,$$

with the maximum attained only at $\eta = \eta_2$. Such a contour Γ exists because $\Xi(1) = 1$, whence, by (3-1) and (3-2), $F_z(\Xi(\eta), \eta) \to 0$ as $\eta \to 1$, and $F_z(\Xi(\eta), \eta) \to \infty$ as $\eta \to z$ or $\eta \to \infty$. Furthermore, for any fixed $\eta \in \Gamma$, the function $F_z(\xi, \eta)$ vanishes at $\xi = 1$ and tends to infinity as $\xi \to \eta$ or $\xi \to \infty$. Hence we can change the integration path $|\xi - \eta| = \varrho_2$ to a closed contour δ_η of steepest descent for $|F_z(\xi, \eta)|$, enclosing η and passing through 1 and through the saddle-point $\xi = \Xi(\eta)$. Thus

$$\max_{\xi \in \delta_\eta} |F_z(\xi, \eta)| = |F_z(\Xi(\eta), \eta)|,$$

with the maximum attained only at $\xi = \Xi(\eta)$. Therefore, by Hata's theorem [2000, Section 1],

$$\lim_{n\to\infty} \frac{1}{n} \log |I_z^{(2)}(hn, jn, kn, ln, mn)| = -(j+k)\log z + \log |F_z(\xi_2, \eta_2)|. \tag{3-13}$$

For the integrals (3-5) and (3-6) we can dispense with the saddle-point method, and apply instead the asymptotic formulae [RV05, (5.16) and (5.19)]; i.e., by (3-3) and (3-12),

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{n} \log I_z^{(0)}(hn, jn, kn, ln, mn) &= -(l+m)\log z + \log f_z(x_0, y_0) \\ &= -(j+k)\log z + \log |F_z(\xi_0, \eta_0)| \end{aligned} \tag{3-14}$$

and

$$\begin{aligned} \limsup_{n\to\infty} \frac{1}{n} \log |I_z^{(1)}(hn, jn, kn, ln, mn)| &\le -(l+m)\log z + \log |f_z(x_1, y_1)| \\ &= -(j+k)\log z + \log |F_z(\xi_1, \eta_1)|. \end{aligned} \tag{3-15}$$

3.2. ***The case $z < 0$.*** Since the polynomial $U(\xi)$ in (3-10) has the leading coefficient

$$(m+h-k)(h+j-l) > 0$$

and takes the value $-jmz > 0$ at $\xi = 0$, the cubic equation (3-10) has a negative root, which we now denote by

$$\xi_2 < 0.$$

Taking into account that

$$U\left(\frac{j}{h+j}\right) = \frac{h^2 jl(j+k-m)}{(h+j)^3} > 0,$$

we choose the integers h, j, k, l, m such that the remaining roots ξ_0 and ξ_1 of $U(\xi)$ are distinct, and either real with

$$\frac{j}{h+j} < \xi_0 < \xi_1 < \frac{j}{m+h-k} < \xi_+, \tag{3-16}$$

or complex conjugate,

$$\xi_1 = \overline{\xi_0} \quad \text{with } \xi_1 \in \mathcal{C}_2, \ \ \xi_0 \in \mathcal{D}_2. \tag{3-17}$$

We first apply the saddle-point method to the integral (3-7). From $\xi_2 < 0$ and (3-11) we get $\xi_2 < \eta_2 < 0$. Again with notation as in [Viola and Zudilin 2018, Sections 4.1 and 4.2], we choose

$$\Delta = \mathcal{C} \cup \mathcal{D} \cup (-\infty, \eta_-), \quad \Xi : \Delta \to \mathcal{C}_1 \cup \mathcal{D}_1 \cup (-\infty, \xi_-), \quad (\xi_*, \eta_*) = (\xi_2, \eta_2).$$

For $\eta \in \Delta$, $\eta \to \infty$, we have $\Xi(\eta) \to \infty$. Thus from (3-2) and (3-9) we see that $F_z(\Xi(\eta), \eta) \to 0$ as $\eta \to 0$, and $F_z(\Xi(\eta), \eta) \to \infty$ as $\eta \to z$ or $\eta \to \infty$. Since $\mathrm{d}F_z(\Xi(\eta), \eta)/\mathrm{d}\eta = 0$ at $\eta = \eta_2$, we obtain

$$\xi_2 < \eta_2 < z < 0.$$

Thus there exists a closed contour $\Gamma \subset \Delta$ of steepest descent for $|F_z(\Xi(\eta), \eta)|$ enclosing z and passing through 0 and η_2. Therefore

$$\max_{\eta \in \Gamma} |F_z(\Xi(\eta), \eta)| = |F_z(\xi_2, \eta_2)|,$$

with the maximum attained only at $\eta = \eta_2$. For any fixed $\eta \in \Gamma$ we have $F_z(0, \eta) = 0$ and $F_z(\xi, \eta) \to \infty$ as $\xi \to \eta$ or $\xi \to \infty$. Hence we change the path $|\xi - \eta| = \varrho_2$ for ξ to a closed contour δ'_η of steepest descent for $|F_z(\xi, \eta)|$, enclosing η and passing through 0 and $\Xi(\eta)$. Thus with the same discussion as in Section 3.1 we get the analogue of (3-13); i.e.,

$$\lim_{n\to\infty} \frac{1}{n} \log |I_z^{(2)}(hn, jn, kn, ln, mn)| = -(j+k)\log|z| + \log|F_z(\xi_2, \eta_2)|. \tag{3-18}$$

For the application of the saddle-point method to the integrals (3-5) and (3-6) we choose

$$\Delta = \mathbb{C} \setminus [\eta_-, \eta_+], \quad \Xi : \Delta \to \mathcal{C}_2 \cup \mathcal{D}_2 \cup (\xi_-, \xi_+),$$

and we distinguish two cases, according to whether (3-16) or (3-17) holds.

First case: $\xi_0, \xi_1 \in \mathbb{R}$ satisfy (3-16). By (3-11) and (3-16) we get

$$\eta_+ < 1 < \eta_1 < \eta_0.$$

For $\eta \in \Delta$ we now have, by (3-9), $\Xi(\eta) \to j/(h+j)$ as $\eta \to \infty$, and $\Xi(\eta) \to (m-k)/(m+h-k)$ as $\eta \to 0$. Hence, by (3-2), $F_z(\Xi(\eta), \eta) \to \infty$ as $\eta \to z$, and $F_z(\Xi(\eta), \eta) \to 0$ as $\eta \to 0$ or $\eta \to \infty$. Since η_0 and η_1 are, respectively, a local maximum and a local minimum of $|F_z(\Xi(\eta), \eta)|$ for $\eta \in \mathbb{R}$ along the half-line $(\eta_+, +\infty)$, the steepest descent direction for $|F_z(\Xi(\eta), \eta)|$ is the direction of the real line at η_0, and the direction orthogonal to the real line at η_1. Hence, by steepest descent, there exists a path γ from 0 to η_1 in the half-plane $\operatorname{Im}\eta < 0$, with tangent at η_1 orthogonal to the real line, such that the maximum of $|F_z(\Xi(\eta), \eta)|$ on γ is attained only at $\eta = \eta_1$. Thus for the integral (3-6) we may take $\Gamma_{0,1} = \gamma \cup \bar{\gamma}$, whence

$$\max_{\eta \in \Gamma_{0,1}} |F_z(\Xi(\eta), \eta)| = |F_z(\xi_1, \eta_1)|$$

with the maximum attained only at $\eta = \eta_1$; accordingly we choose $(\xi_*, \eta_*) = (\xi_1, \eta_1)$. For any fixed $\eta \in \Gamma_{0,1}$, $\eta \neq 0$, by (3-2) we have $F_z(0, \eta) = F_z(1, \eta) = 0$ and $F_z(\xi, \eta) \to \infty$ as $\xi \to \eta$ or $\xi \to \infty$. Since $\operatorname{Im}\eta$ and $\operatorname{Im}\Xi(\eta)$ have opposite signs, keeping the endpoints $\xi = 0$ and $\xi = 1$ fixed we can continuously deform the integration interval $[0, 1]$ for ξ, without encountering η, to an integration path δ''_η of steepest descent for $|F_z(\xi, \eta)|$, equivalent to $[0, 1]$ by Cauchy's theorem and passing through the saddle-point $\xi = \Xi(\eta)$. Hence

$$\max_{\xi \in \delta''_\eta} |F_z(\xi, \eta)| = |F_z(\Xi(\eta), \eta)| \tag{3-19}$$

with the maximum attained only at $\xi = \Xi(\eta)$. By Hata's theorem,

$$\lim_{n\to\infty} \frac{1}{n} \log |I_z^{(1)}(hn, jn, kn, ln, mn)| = -(j+k)\log|z| + \log|F_z(\xi_1, \eta_1)|. \tag{3-20}$$

For the integral (3-5) we choose $\operatorname{Im}\zeta < 0$. Again by the argument in (2-6), the integration half-line $(0, \zeta\infty)$ for η can be transformed to $\gamma \cup [\eta_1, +\infty)$, where γ is defined as above, without changing the value of (3-5). Clearly

$$\max_{\eta \in \gamma \cup [\eta_1, +\infty)} |F_z(\Xi(\eta), \eta)| = |F_z(\xi_0, \eta_0)|,$$

with the maximum attained only at $\eta = \eta_0$. Thus we take here $(\xi_*, \eta_*) = (\xi_0, \eta_0)$. For any fixed $\eta \in \gamma \cup [\eta_1, +\infty)$ we have $\Xi(\eta) \in \mathcal{D}_2 \cup (\xi_-, \xi_+)$. Hence, as above, the integration interval $[0, 1]$ for ξ in (3-5) can be deformed to a path δ''_η from 0 to 1 passing through $\Xi(\eta)$, equivalent to $[0, 1]$ by Cauchy's theorem and satisfying (3-19). Therefore

$$\lim_{n\to\infty} \frac{1}{n} \log |I_z^{(0)}(hn, jn, kn, ln, mn)| = -(j+k)\log|z| + \log|F_z(\xi_0, \eta_0)|. \tag{3-21}$$

Second case: $\xi_1 \in \mathcal{C}_2$, $\xi_0 = \overline{\xi_1} \in \mathcal{D}_2$. Now $\eta_1 \in \mathcal{C}$, $\eta_0 = \overline{\eta_1} \in \mathcal{D}$. Since $F_z(\Xi(\eta), \eta) \to 0$ as $\eta \to 0$ or $\eta \to \infty$, there exists a path $\gamma' \subset \mathcal{D}$ from 0 to a sufficiently large $\eta' > 0$, passing through the saddle-point η_0, such that

$$\max_{\eta \in \gamma'} |F_z(\Xi(\eta), \eta)| = |F_z(\xi_0, \eta_0)|$$

with the maximum attained only at η_0. For the integral (3-6) we take $\Gamma_{0,1} = \gamma' \cup \overline{\gamma'}$, whence

$$\max_{\eta \in \Gamma_{0,1}} |F_z(\Xi(\eta), \eta)| = |F_z(\xi_0, \eta_0)| = |F_z(\xi_1, \eta_1)|.$$

Thus we must apply the saddle-point method separately for $\eta \in \gamma'$ and $\eta \in \overline{\gamma'}$. We have

$$\left|\oint_{\Gamma_{0,1}} \left(\int_0^1 F_z(\xi,\eta)^n \frac{\mathrm{d}\xi}{\xi-\eta}\right)\frac{\mathrm{d}\eta}{\eta-z}\right| \le \left|\int_{\gamma'} \left(\int_0^1 F_z(\xi,\eta)^n \frac{\mathrm{d}\xi}{\xi-\eta}\right)\frac{\mathrm{d}\eta}{\eta-z}\right| + \left|\int_{\overline{\gamma'}} \left(\int_0^1 F_z(\xi,\eta)^n \frac{\mathrm{d}\xi}{\xi-\eta}\right)\frac{\mathrm{d}\eta}{\eta-z}\right|. \quad (3\text{-}22)$$

For any fixed $\eta \in \gamma'$ or $\eta \in \overline{\gamma'}$, $\eta \neq 0$, we have $F_z(0,\eta) = F_z(1,\eta) = 0$ and $F_z(\xi,\eta) \to \infty$ as $\xi \to \eta$ or $\xi \to \infty$. As above, there exists an integration path δ''_η for ξ from 0 to 1 passing through $\Xi(\eta)$, equivalent to $[0, 1]$ by Cauchy's theorem and satisfying (3-19). By Hata's theorem [2000, (1.9)], the quotient of the absolute values of the integrals over γ' and $\overline{\gamma'}$ in (3-22) tends to a constant as $n \to \infty$, since $|F_z(\xi_0,\eta_0)| = |F_z(\xi_1,\eta_1)|$. It follows that

$$\limsup_{n\to\infty} \frac{1}{n} \log |I_z^{(1)}(hn, jn, kn, ln, mn)| \le -(j+k)\log|z| + \log|F_z(\xi_1,\eta_1)|. \quad (3\text{-}23)$$

In (3-5) we choose $\operatorname{Im}\zeta < 0$, and we change the integration half-line $(0, \zeta\infty)$ for η to a suitable path $\gamma'' \subset \mathcal{D}$ from 0 to ∞, passing through the saddle-point η_0, so that

$$\max_{\eta\in\gamma''} |F_z(\Xi(\eta),\eta)| = |F_z(\xi_0,\eta_0)|$$

with the maximum attained only at η_0. Again, for any $\eta \in \gamma''$ there exists a path δ''_η for ξ from 0 to 1 passing through $\Xi(\eta)$, equivalent to $[0, 1]$ by Cauchy's theorem and satisfying (3-19). Therefore

$$\lim_{n\to\infty} \frac{1}{n} \log |I_z^{(0)}(hn, jn, kn, ln, mn)| = -(j+k)\log|z| + \log|F_z(\xi_0,\eta_0)|. \quad (3\text{-}24)$$

4. Linear independence measures

In (2-27) we set $z = \frac{s}{r}$, with integers $r \ge 1$ and s satisfying either $s > r$ or $s < 0$, and we multiply by $r^{\delta n}$. We obtain the following extension of [RV05, (4.12)]:

$$\begin{aligned} D_n s^{\alpha n}(s-r)^{\beta n} r^{(\delta-\alpha-\beta)n} I_{s/r}(hn, jn, kn, ln, mn) &= p_n - q_n \mathrm{Li}_2\Big(\frac{r}{s}\Big),\\ D_n s^{\alpha n}(s-r)^{\beta n} r^{(\delta-\alpha-\beta)n} I^{(1)}_{s/r}(hn, jn, kn, ln, mn) &= p'_n - q_n \mathrm{Li}_1\Big(\frac{r}{s}\Big),\\ D_n s^{\alpha n}(s-r)^{\beta n} r^{(\delta-\alpha-\beta)n} I^{(2)}_{s/r}(hn, jn, kn, ln, mn) &= q_n, \end{aligned} \quad (4\text{-}1)$$

with

$$p_n = r^{\delta n} P_n^*\Big(\frac{s}{r}\Big), \quad q_n = r^{\delta n} Q_n^*\Big(\frac{s}{r}\Big), \quad p'_n = r^{\delta n} R_n^*\Big(\frac{s}{r}\Big).$$

By (2-28),

$$p_n, p'_n, q_n \in \mathbb{Z}.$$

Let, by (3-3) and (3-12),

$$\begin{aligned} c_\nu &= -\log|f_{s/r}(x_\nu, y_\nu)| = (j+k-l-m)\log\frac{|s|}{r} - \log|F_{s/r}(\xi_\nu,\eta_\nu)| \quad (\nu = 0, 1),\\ c_2 &= \log|f_{s/r}(x_2, y_2)| = (l+m-j-k)\log\frac{|s|}{r} + \log|F_{s/r}(\xi_2,\eta_2)|, \end{aligned}$$

and

$$c_3 = M + N - \left(\int_{\Omega} \mathrm{d}\psi(x) + \int_{\Omega'} \mathrm{d}\psi(x)\right) + (\alpha - l - m)\log|s| + (m + h - k)\log r + \beta \log|s - r|. \quad (4\text{-}2)$$

Since, by (2-16), $\delta - \alpha - \beta = h - k - l$, applying (2-29), (3-13) and (3-18) in the last equation of (4-1) we obtain

$$\lim_{n\to\infty} \frac{1}{n} \log|q_n| = c_2 + c_3. \quad (4\text{-}3)$$

Similarly, defining

$$r_n^{(1)} = p_n' - q_n \mathrm{Li}_1\Big(\frac{r}{s}\Big), \quad r_n^{(2)} = p_n - q_n \mathrm{Li}_2\Big(\frac{r}{s}\Big),$$

and using (2-29), (3-14), (3-15), (3-20), (3-21), (3-23) and (3-24) in the first two equations of (4-1), we get

$$\limsup_{n\to\infty} \frac{1}{n} \log|r_n^{(1)}| \le c_3 - c_1 \quad (4\text{-}4)$$

and, by (2-4),

$$\limsup_{n\to\infty} \frac{1}{n} \log|r_n^{(2)}| \le c_3 - \min\{c_0, c_1\}. \quad (4\text{-}5)$$

Moreover, if $c_0 < c_1$,

$$\lim_{n\to\infty} \frac{1}{n} \log|r_n^{(2)}| = c_3 - c_0, \quad (4\text{-}6)$$

by virtue of (3-14), (3-21) and (3-24). Thus, for a given $z = \frac{s}{r}$, if we choose the integers $h,\ j,\ k,\ l,\ m$ such that $c_3 < c_0 < c_1$, applying [Viola and Zudilin 2018, Lemma 5.1] with $S = 2,\ \gamma_1 = \mathrm{Li}_1\big(\frac{r}{s}\big),\ \gamma_2 = \mathrm{Li}_2\big(\frac{r}{s}\big)$ and $\tau = c_0 - c_3$, from (4-4) and (4-6) we deduce the linear independence over $\mathbb{Q}$ of 1, $\mathrm{Li}_1\big(\frac{r}{s}\big)$ and $\mathrm{Li}_2\big(\frac{r}{s}\big)$. Once the linear independence is established, in order to improve the $\mathbb{Q}$-linear independence measure of 1, $\mathrm{Li}_1\big(\frac{r}{s}\big)$ and $\mathrm{Li}_2\big(\frac{r}{s}\big)$ that can be obtained through the above choice of $h,\ j,\ k,\ l,\ m$, we can make a new choice of $h,\ j,\ k,\ l,\ m$, possibly different from the previous one, with $c_0 \lesseqgtr c_1$ and $c_3 < \min\{c_0, c_1\}$. Then, by Hata's lemma for $S = 2$ generalized by Marcovecchio for any S, i.e., by [Viola and Zudilin 2018, Lemma 5.2] with $S = 2,\ \gamma_1 = \mathrm{Li}_1\big(\frac{r}{s}\big),\ \gamma_2 = \mathrm{Li}_2\big(\frac{r}{s}\big),\ d = 0,\ \sigma = c_2 + c_3$ and $\tau = \min\{c_0, c_1\} - c_3$, from (4-3), (4-4) and (4-5) we get for 1, $\mathrm{Li}_1\big(\frac{r}{s}\big)$ and $\mathrm{Li}_2\big(\frac{r}{s}\big)$ the $\mathbb{Q}$-linear independence measure

$$\frac{c_2 + c_3}{c - c_3} \quad \text{with } c = \min\{c_0, c_1\}; \quad (4\text{-}7)$$

i.e., for any $\varepsilon > 0$ and any $(a_0, a_1, a_2) \in \mathbb{Z}^3 \setminus \{(0, 0, 0)\}$,

$$\left|a_0 + a_1 \mathrm{Li}_1\Big(\frac{r}{s}\Big) + a_2 \mathrm{Li}_2\Big(\frac{r}{s}\Big)\right| > C(\varepsilon) A^{-(c_2+c_3)/(c-c_3)-\varepsilon},$$

where $A = \max\{|a_0|, |a_1|, |a_2|\}$ and where the constant $C(\varepsilon) > 0$ is independent of (a_0, a_1, a_2).

We separately treat the positive and the negative cases, $z > 1$ and $z < 0$. From [RV05, Theorem 5.2] where $c_3 < c_0 < c_1$ is assumed, and from the subsequent discussion in Section 6 of that paper, we obtain the $\mathbb{Q}$-linear independence of 1, $\mathrm{Li}_1\big(\frac{r}{s}\big)$ and $\mathrm{Li}_2\big(\frac{r}{s}\big)$ for all integers r, s with $r \ge 1$ and $s \ge s_+(r) > r$, where $s_+(r)$ is explicit. In [RV05] we found the values $s_+(1) = 6,\ s_+(2) = 51,\ s_+(3) = 173,\ s_+(4) = 423$. Concerning $\mathbb{Q}$-linear independence measures, applying [Viola and Zudilin 2018, Lemma 5.2] we get a new proof of the linear independence measures of 1, $\mathrm{Li}_1\big(\frac{1}{z}\big)$ and $\mathrm{Li}_2\big(\frac{1}{z}\big)$ for $z = \frac{s}{r} > 1$ given in [Marcovecchio 2016, Table p. 231], by choosing $h = \gamma,\ j = \alpha_1,\ k = \beta_1 + \alpha_2 - \alpha_1,\ l = \beta_2 + \alpha_1 - \alpha_2$,

	z	h	j	k	l	m	$(c_2+c_3)/(c-c_3)$
$*$	-5	10	8	7	6	10	61.68698…
$*$	-6	10	8	7	6	10	43.51295…
$*$	-7	64	32	37	27	43	32.76353…
$*$	-8	64	32	37	27	43	25.05713…
$*$	-9	64	32	37	27	43	20.87789…
$*$	-10	64	32	37	27	43	18.24158…
$*$	-11	64	32	37	27	43	16.41900…
	-12	64	32	37	27	43	15.08001…
$*$	-13	130	65	76	54	86	14.04958…
	-14	130	65	76	54	86	13.22908…
	-15	130	65	76	54	86	12.58420…
	-16	130	65	76	54	86	12.05561…
	-17	130	65	76	54	86	11.61188…
	-18	130	65	76	54	86	11.23276…
	-19	130	65	76	54	86	10.90426…
	-20	130	65	76	54	86	10.61629…

Table 1

$m = \alpha_2$ for each z, where α_1, β_1, α_2, β_2, γ are the integers in Marcovecchio's table. Note, however, that in Marcovecchio's notation the quantity μ appearing in the last column of his table equals our linear independence measure (4-7) plus 1; i.e.,

$$\mu = \frac{c_2 + c_3}{c - c_3} + 1 = \frac{c + c_2}{c - c_3}.$$

In the negative case, our method yields new $\mathbb{Q}$-linear independence measures of 1, $\mathrm{Li}_1\left(\frac{r}{s}\right)$ and $\mathrm{Li}_2\left(\frac{r}{s}\right)$ for integers $r \geq 1$ and $s \leq s_-(r) < 0$ with explicit $s_-(r)$, improving both the results in [Hata 1993, Table 2 p. 386] and in [Marcovecchio 2016, Table p. 231]. For brevity we give numerical results only for $r = 1$, and for $-20 \leq z = s \leq s_-(1) = -5$. For such values of z it is easy to prove the $\mathbb{Q}$-linear independence of 1, $\mathrm{Li}_1\left(\frac{1}{z}\right)$ and $\mathrm{Li}_2\left(\frac{1}{z}\right)$, e.g., with the choice $h = 20,\ j = 10,\ k = 13,\ l = 8,\ m = 16$, which yields $c_3 < c_0 < c_1$ for all $z \in \mathbb{Z}$ satisfying $-20 \leq z \leq -5$, and hence allows one to apply [Viola and Zudilin 2018, Lemma 5.1]. Then for each z we apply Lemma 5.2 of the same paper with the corresponding values of $h,\ j,\ k,\ l,\ m$ in Table 1, thus obtaining for 1, $\mathrm{Li}_1\left(\frac{1}{z}\right)$ and $\mathrm{Li}_2\left(\frac{1}{z}\right)$ the $\mathbb{Q}$-linear independence measure $(c_2 + c_3)/(c - c_3)$ in the last column of the table.

The values of z in Table 1 marked with an asterisk are associated with $h,\ j,\ k,\ l,\ m$ for which (3-17) holds, i.e., such that the polynomial $U(\xi)$ in (3-10) has complex conjugate roots ξ_0, ξ_1, whence $c = c_0 = c_1$. For the remaining values of z, the corresponding $h,\ j,\ k,\ l,\ m$ yield (3-16) with $c = c_0 < c_1$.

As an example we give below all the numerical values for $z = -5$, with $h = 10,\ j = 8,\ k = 7,\ l = 6$, $m = 10$. We get

$$l + m - j = 8, \quad m + h - k = 13, \quad h + j - l = 12, \quad j + k - m = 5,$$

whence

$$M = 13, \quad N = 12, \quad \alpha = l + m = 16, \quad \beta = k + l - h = 3,$$

$$\Omega = \left[\tfrac{1}{10}, \tfrac{1}{4}\right) \cup \left[\tfrac{3}{10}, \tfrac{5}{13}\right) \cup \left[\tfrac{2}{5}, \tfrac{5}{12}\right) \cup \left[\tfrac{1}{2}, \tfrac{7}{13}\right) \cup \left[\tfrac{4}{7}, \tfrac{5}{8}\right) \cup \left[\tfrac{7}{10}, \tfrac{10}{13}\right) \cup \left[\tfrac{4}{5}, \tfrac{11}{13}\right) \cup \left[\tfrac{6}{7}, \tfrac{7}{8}\right) \cup \left[\tfrac{9}{10}, \tfrac{12}{13}\right),$$

$$\int_\Omega \mathrm{d}\psi(x) = 7.87642\ldots, \quad \Omega' = \varnothing,$$

$$c_3 = M + N - \int_\Omega \mathrm{d}\psi(x) + \beta \log 6 = 22.49885\ldots.$$

The saddle-points of $F_{-5}(\xi, \eta)$ are

$$(\xi_0, \eta_0) = (0.45905\ldots + i\,0.04354\ldots,\ \ 0.96082\ldots - i\,1.38422\ldots),$$
$$(\xi_1, \eta_1) = (\overline{\xi_0}, \overline{\eta_0}),$$
$$(\xi_2, \eta_2) = (-12.05913\ldots,\ \ -8.56053\ldots).$$

Therefore

$$c = c_0 = c_1 = 23.60655\ldots, \quad c_2 = 45.83193\ldots,$$

whence we obtain the $\mathbb{Q}$-linear independence measure of 1, $\mathrm{Li}_1\left(-\frac{1}{5}\right)$, $\mathrm{Li}_2\left(-\frac{1}{5}\right)$ given by

$$\frac{c_2 + c_3}{c - c_3} = 61.68698\ldots.$$

References

[Hata 1993] M. Hata, "Rational approximations to the dilogarithm", *Trans. Amer. Math. Soc.* **336**:1 (1993), 363–387. MR Zbl

[Hata 2000] M. Hata, "$\mathbb{C}^2$-saddle method and Beukers' integral", *Trans. Amer. Math. Soc.* **352**:10 (2000), 4557–4583. MR Zbl

[Marcovecchio 2016] R. Marcovecchio, "Linear independence of polylogarithms at algebraic points", *Mosc. J. Comb. Number Theory* **6**:2-3 (2016), 208–232. MR Zbl

[Rhin and Viola 1996] G. Rhin and C. Viola, "On a permutation group related to $\zeta(2)$", *Acta Arith.* **77**:1 (1996), 23–56. MR Zbl

[Rhin and Viola 2005] G. Rhin and C. Viola, "The permutation group method for the dilogarithm", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **4**:3 (2005), 389–437. MR Zbl

[Viola and Zudilin 2018] C. Viola and W. Zudilin, "Linear independence of dilogarithmic values", *J. Reine Angew. Math.* **736** (2018), 193–223. MR Zbl

Received 10 Jan 2018.

GEORGES RHIN:

georges.rhin@univ-lorraine.fr
IECL, Université de Lorraine, UFR MIM, Metz, France

CARLO VIOLA:

viola@dm.unipi.it
Dipartimento di Matematica, Università di Pisa, Pisa, Italy

Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the submission page.

Originality. Submission of a manuscript acknowledges that the manuscript is original and and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

Language. Articles are usually in English or French, but articles written in other languages are welcome.

Required items. A brief abstract of about 150 words or less must be included. It should be self-contained and not refer to bibliography keys. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and a Mathematics Subject Classification for the article, and, for each author, affiliation (if appropriate) and email address.

Format. Authors are encouraged to use LaTeX and the standard amsart class, but submissions in other varieties of TeX, and exceptionally in other formats, are acceptable. Initial uploads should normally be in PDF format; after the refereeing process we will ask you to submit all source material.

References. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibTeX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

Figures. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages — Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc. — allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with as many details as you can about how your graphics were generated.

Bundle your figure files into a single archive (using zip, tar, rar or other format of your choice) and upload on the link you been provided at acceptance time. Each figure should be captioned and numbered so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text ("the curve looks like this:"). It is acceptable to submit a manuscript with all figures at the end, if their placement is specified in the text by means of comments such as "Place Figure 1 here". The same considerations apply to tables.

White Space. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Proofs. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

Moscow Journal of Combinatorics and Number Theory
vol. 8 no. 1 2019