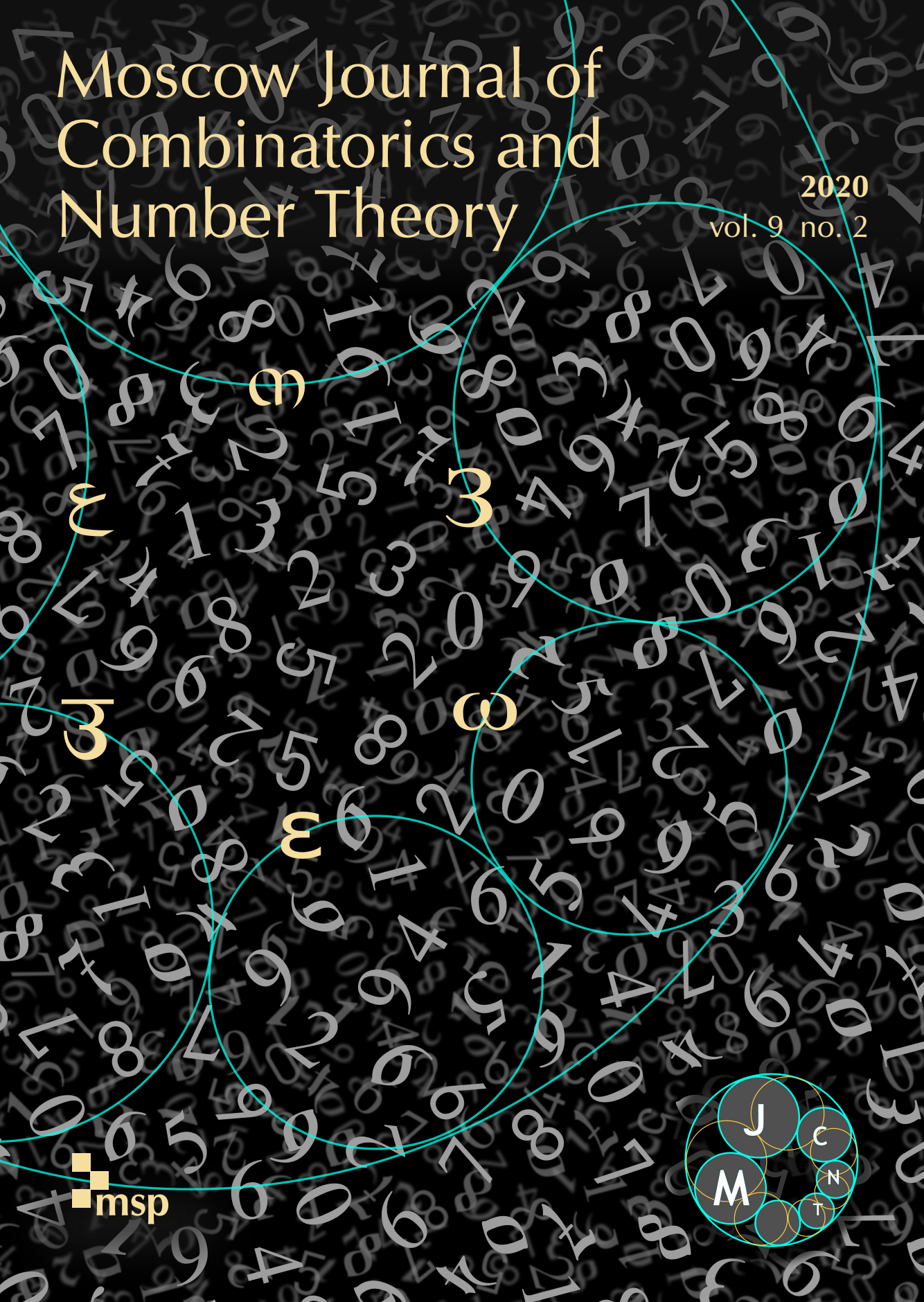
# Moscow Journal of Combinatorics and Number Theory

2020
vol. 9 no. 2



msp

# Moscow Journal of Combinatorics and Number Theory

msp.org/moscow





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



The subscription price for 2020 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 2640-7361 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/

msp

Moscow Journal of Combinatorics and Number Theory
Vol. 9, No. 2, 2020
https://doi.org/10.2140/moscow.2020.9.101

# A dynamical Borel–Cantelli lemma via improvements to Dirichlet's theorem

Dmitry Kleinbock and Shucheng Yu

Let $X \cong \mathrm{SL}_2(\mathbb{R})/\mathrm{SL}_2(\mathbb{Z})$ be the space of unimodular lattices in $\mathbb{R}^2$, and for any $r \geq 0$ denote by $K_r \subset X$ the set of lattices such that all its nonzero vectors have supremum norm at least $e^{-r}$. These are compact nested subsets of $X$, with $K_0 = \bigcap_r K_r$ being the union of two closed horocycles. We use an explicit second moment formula for the Siegel transform of the indicator functions of squares in $\mathbb{R}^2$ centered at the origin to derive an asymptotic formula for the volume of sets $K_r$ as $r \to 0$. Combined with a zero-one law for the set of the $\psi$-Dirichlet numbers established by Kleinbock and Wadleigh (*Proc. Amer. Math. Soc.* **146** (2018), 1833–1844), this gives a new dynamical Borel–Cantelli lemma for the geodesic flow on $X$ with respect to the family of shrinking targets $\{K_r\}$.

## 1. Introduction

Let $(X, \mu)$ be a probability space, and let $\{a_s\}_{s\in\mathbb{R}}$ be a one-parameter measure-preserving flow on $X$. Given a family of measurable subsets $\{B_s\}_{s>0}$ of $X$ with $\mu(B_s) \to 0$ as $s \to \infty$ (called *shrinking targets*), the *shrinking targets problem* asks for a dichotomy on whether generic orbits of $\{a_s\}_{s>0}$ would hit the shrinking targets indefinitely. That is, we are looking for a zero-one law for the measure of the limsup set

$$B_\infty := \limsup_{s\to\infty} a_{-s} B_s = \{x \in X \mid a_s x \in B_s \text{ for an unbounded set of } s > 0\}.$$

For any $n \in \mathbb{N}$ let

$$\widetilde{B}_n := \bigcup_{0\leq s<1} a_{-s} B_{n+s} \tag{1-1}$$

be the thickening of the shrinking targets $\{B_s\}_{n\leq s<n+1}$ along the flow $\{a_{-s}\}_{0\leq s<1}$. Note that $a_n x \in \widetilde{B}_n$ if and only if there exists some $s \in [n, n+1)$ such that $a_s x \in B_s$. We thus have

$$B_\infty = \limsup_{n\to\infty} a_{-n} \widetilde{B}_n = \{x \in X \mid a_n x \in \widetilde{B}_n \text{ infinitely often}\}, \tag{1-2}$$

and the classical Borel–Cantelli lemma implies

$$\sum_n \mu(\widetilde{B}_n) < \infty \quad \Longrightarrow \quad \mu(B_\infty) = 0. \tag{1-3}$$

Kleinbock was supported by NSF grants DMS-1600814 and DMS-1900560. Yu acknowledges that this project received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement no. 754475).

*MSC2010:* primary 11J04, 37A17; secondary 11H60, 37D40.

*Keywords:* Siegel transform, dynamical Borel–Cantelli lemma.

On the other hand, following the terminology of [Chernov and Kleinbock 2001] we say the family of shrinking targets $\{B_s\}_{s>0}$ is *Borel–Cantelli* (*BC*) for the flow $\{a_s\}_{s>0}$ if $\mu(B_\infty) = 1$. Thus a necessary condition for $\{B_s\}_{s>0}$ to be BC for $\{a_s\}_{s>0}$ is that the sequence of its thickenings has divergent sum of measures, and we say $\{B_s\}_{s>0}$ satisfies *a dynamical Borel–Cantelli lemma for* $\{a_s\}_{s>0}$ if this is also a sufficient condition.

The shrinking targets problem for continuous time flow in the context of homogeneous spaces was first studied in [Sullivan 1982], where he established a logarithm law for the fastest rate of geodesic cusp excursions in finite-volume hyperbolic manifolds. Later using the exponential mixing rate and a smooth approximation argument, the first author and Margulis [Kleinbock and Margulis 1999] proved that the family of cusp neighborhoods $\{\Phi^{-1}(r(s), \infty)\}_{s>0}$ with divergent sum of measures is BC for any diagonalizable flow on $(G/\Gamma, \mu)$, where $G$ is a connected semisimple Lie group without compact factors, $\Gamma < G$ is an irreducible lattice, and $\mu$ is the probability measure on $X = G/\Gamma$ coming from a Haar measure on $G$. Here $\Phi$ is a distance-like function on $X$ [loc. cit., Definition 1.6] and $r(\cdot)$ is a quasi-increasing function [loc. cit., Section 2.4]. Later Maucourant [2006] obtained a similar dynamical Borel–Cantelli lemma for geodesic flows making excursions into shrinking hyperbolic balls (with a fixed center) on a finite-volume hyperbolic manifold. See [Athreya 2009] for a survey on shrinking targets problems in dynamical systems.

One main reason that such dynamical Borel–Cantelli lemmas have gained much attention is due to their connections to metric number theory, which were first explored in [Sullivan 1982]. Such connections were made more apparent later in [Kleinbock and Margulis 1999]. Let $m, l$ be two positive integers and let $M_{m,l}(\mathbb{R})$ be the space of $m$ by $l$ real matrices. Given $\psi : [t_0, \infty) \to (0, \infty)$ a continuous nonincreasing function, let us define $W(\psi) \subset M_{m,l}(\mathbb{R})$, the set of $\psi$-approximable $m \times l$ real matrices such that $A \in W(\psi)$ if and only if there are infinitely many $\boldsymbol{q} \in \mathbb{Z}^l$ satisfying

$$\|A\boldsymbol{q} - \boldsymbol{p}\|^m < \psi(\|\boldsymbol{q}\|^l) \quad \text{for some } \boldsymbol{p} \in \mathbb{Z}^m,$$

where $\|\cdot\|$ is the supremum norm on respective Euclidean spaces. The classical Khinchin–Groshev theorem gives an exact criterion on when $W(\psi)$ has full or zero Lebesgue measure.

**Theorem KG** (Khinchin–Groshev). *Given a continuous nonincreasing $\psi$, the set $W(\psi)$ has full (resp. zero) Lebesgue measure if and only if the series $\sum_k \psi(k)$ diverges (resp. converges).*

See [Schmidt 1980] for more details. On the other hand, let $X = \mathrm{SL}_{m+l}(\mathbb{R})/\mathrm{SL}_{m+l}(\mathbb{Z})$ be the space of unimodular lattices in $\mathbb{R}^{m+l}$ and let $\Delta : X \to [0, \infty)$ be the function on $X$ given by

$$\Delta(\Lambda) := \sup_{\boldsymbol{v} \in \Lambda \smallsetminus \{\boldsymbol{0}\}} \log\left(\frac{1}{\|\boldsymbol{v}\|}\right). \tag{1-4}$$

Note that $\Delta(\Lambda) \geq 0$ for any $\Lambda \in X$ due to Minkowski's convex body theorem, and for all $r \geq 0$ the sets

$$K_r := \Delta^{-1}([0, r]) \tag{1-5}$$

(of lattices such that all its nonzero vectors have supremum norm at least $e^{-r}$) are compact due to Mahler's compactness criterion; see, e.g., [Cassels 1997]. Following ideas of [Dani 1985], it was shown in [Kleinbock and Margulis 1999] that there exists a unique function $r = r_\psi : [s_0, \infty) \to \mathbb{R}$ depending

on $\psi$ (this was referred to as the Dani correspondence) such that $A \in M_{m,l}(\mathbb{R})$ is $\psi$-approximable if and only if the events $a_s \Lambda_A \in \Delta^{-1}(r(s), \infty)$ happen for an unbounded set of $s > s_0$, where

$$a_s = \operatorname{diag}(e^{s/m}, \dots, e^{s/m}, e^{-s/l}, \dots, e^{-s/l}),$$

with $m$ copies of $e^{s/m}$ and $l$ copies of $e^{-s/l}$, and

$$\Lambda_A = \begin{pmatrix} I_m & A \\ 0 & I_l \end{pmatrix} \mathbb{Z}^{m+l} \in X.$$

This way the first author and Margulis showed Theorem KG to be equivalent to a dynamical Borel–Cantelli lemma for the $a_s$-orbits making excursions into the cusp neighborhoods $\Delta^{-1}(r(s), \infty)_{s>s_0}$, and used this to give an alternative dynamical proof of Theorem KG based on mixing properties of the $a_s$-action on $X$; see [Kleinbock and Margulis 1999].

More recently, for a given $\psi$ as above, the first author and Wadleigh [Kleinbock and Wadleigh 2018] studied the finer problem of improvements to Dirichlet's theorem. See [Davenport and Schmidt 1970a; 1970b] for the history of the problem of improving Dirichlet's theorem. Following the definition in [Kleinbock and Wadleigh 2018] an $m$ by $l$ real matrix $A$ is called *$\psi$-Dirichlet* if the system of inequalities

$$\|A\boldsymbol{q} - \boldsymbol{p}\|^m < \psi(t) \quad \text{and} \quad \|\boldsymbol{q}\|^l < t$$

has solutions in $(\boldsymbol{p}, \boldsymbol{q}) \in \mathbb{Z}^m \times (\mathbb{Z}^l \smallsetminus \{\boldsymbol{0}\})$ for all sufficiently large $t$. Following the general scheme developed in [Kleinbock and Margulis 1999] they gave a dynamical interpretation of $\psi$-Dirichlet matrices. Namely, they showed that $A \in M_{m,l}(\mathbb{R})$ is not $\psi$-Dirichlet if and only if the events

$$a_s \Lambda_A \in K_{r(s)}$$

happen for an unbounded set of $s > s_0$, where $a_s$, $\Lambda_A$ and $r = r_\psi$ are all as above. Hence in this case the family of shrinking targets is given by $\{K_{r(s)}\}_{s>s_0}$, and one is naturally interested in whether this family of shrinking targets is BC for the flow $\{a_s\}_{s>0}$.

However this dynamical interpretation is not helpful when it comes to determining necessary and sufficient conditions on $\psi$ guaranteeing that almost every (almost no) $A$ is $\psi$-Dirichlet. One of the main difficulties is that the shrinking targets $K_{r(s)}$ are far away from being $\mathrm{SO}_{m+l}(\mathbb{R})$-invariant, and thus applying the mixing properties of the $a_s$-action will involve certain Sobolev norms which are hard to control. Still, using a different method based on continued fractions the aforementioned conditions were found in [Kleinbock and Wadleigh 2018] for the case $m = l = 1$. Namely, the following was proved:

**Theorem KW** (Kleinbock–Wadleigh). *Let $\psi : [t_0, \infty) \to (0, \infty)$ be a continuous, nonincreasing function satisfying*

$$\text{the function } t \mapsto t\psi(t) \text{ is nondecreasing} \tag{1-6}$$

*and*

$$t\psi(t) < 1 \quad \text{for all } t \geq t_0. \tag{1-7}$$

*Then if the series*

$$\sum_n \frac{-(1 - n\psi(n)) \log(1 - n\psi(n))}{n} \tag{1-8}$$

*diverges (resp. converges), then Lebesgue-a.e. $x \in \mathbb{R}$ is not (resp. is) $\psi$-Dirichlet.*

In this paper we use the above theorem to derive a dynamical Borel–Cantelli lemma for the diagonal flow $a_s := \mathrm{diag}(e^s, e^{-s})$ on $X := \mathrm{SL}_2(\mathbb{R})/\mathrm{SL}_2(\mathbb{Z})$. Let $\mu$ be the probability Haar measure on $X$, consider the function $\Delta$ on $X$ as in (1-4), and define the sets $K_r$ as in (1-5).

We now state our dynamical Borel–Cantelli lemma.

**Theorem 1.1.** *Let $r : [s_0, \infty) \to (0, \infty)$ be a continuous and nonincreasing function. Let $B_s = K_{r(s)}$ and let $B_\infty = \limsup_{t\to\infty} a_{-s} B_s$. Then we have*

$$\sum_n r(n) \log\left(\frac{1}{r(n)}\right) < \infty \quad \Longrightarrow \quad \mu(B_\infty) = 0.$$

*If in addition we assume that the function $s \mapsto s + r(s)$ is nondecreasing, then we have*

$$\sum_n r(n) \log\left(\frac{1}{r(n)}\right) = \infty \quad \Longrightarrow \quad \mu(B_\infty) = 1.$$

Comparing the statement of the above theorem with (1-3), one can guess that it can be approached by studying the thickenings

$$\widetilde{B}_n = \bigcup_{0\le s<1} a_{-s} B_{n+s} = \bigcup_{0\le s<1} a_{-s} K_{r(n+s)} \tag{1-9}$$

as in (1-1). We do it in several steps. In the beginning of Section 3 we prove an asymptotic measure formula for the sets $K_r$ where $r$ is small:

**Theorem 1.2.** *For any $0 < r < (\log 2)/2$ we have*

$$\mu(K_r) = \frac{4r^2 \log(1/r)}{\zeta(2)} + O(r^2),$$

*where $\zeta(2) = \pi^2/6$ is the value of the Riemann zeta function at 2.*

Here and hereafter for two positive quantities $A$ and $B$, we will use the notation $A \ll B$ or $A = O(B)$ to mean that there is a constant $c > 0$ such that $A \le cB$, and we will use subscripts to indicate the dependence of the constant on parameters. We will write $A \asymp B$ for $A \ll B \ll A$.

The next step is to use Theorem 1.2 to estimate the measure of the thickening of $K_r$ along the flow $\{a_{-s}\}_{0\le s<1}$ by bounding it from above and below by a finite union of $a_s$-translates of $K_r$. This is also done in Section 3 and yields the following result:

**Theorem 1.3.** *For any $0 < r < \log 1.01$ we have*

$$\mu\left(\bigcup_{0\le s<1} a_{-s} K_r\right) \asymp r \log\left(\frac{1}{r}\right).$$

The above asymptotic equality shows that the series appearing in Theorem 1.1 converges/diverges if and only if so does the series $\sum_n \mu(\widetilde{B}_n)$, where $\widetilde{B}_n$ is as in (1-9):

**Corollary 1.4.** *Let $r : [s_0, \infty) \to (0, \infty)$ be a nonincreasing function, and let $\widetilde{B}_n$ be as in (1-9). Then we have*

$$\sum_n \mu(\widetilde{B}_n) = \infty \quad \Longleftrightarrow \quad \sum_n r(n) \log\left(\frac{1}{r(n)}\right) = \infty.$$

Therefore, in view of (1-2) and (1-3), the convergence part of Theorem 1.1 is immediate from the Borel–Cantelli lemma. The divergence part however is trickier. Instead of using a dynamical approach as in [Kleinbock and Margulis 1999], our proof in Section 4 is non-dynamical and relies on Theorem KW and the Dani correspondence.

It remains to comment on our proof of Theorem 1.2. Instead of trying to describe the sets $K_r$ explicitly in terms of coordinates and compute their measures directly, we adapt an indirect approach which relies on an explicit second moment formula of the Siegel transform of certain indicator functions. Recall that if $f$ is a function on $\mathbb{R}^2$, its *primitive Siegel transform* is the function on $X$ given by

$$\hat{f}(\Lambda) := \sum_{\boldsymbol{v} \in \Lambda_{\mathrm{pr}}} f(\boldsymbol{v}),$$

where $\Lambda_{\mathrm{pr}}$ is the set of primitive vectors of $\Lambda$. Clearly $\hat{f}(\Lambda) = \#(\Lambda_{\mathrm{pr}} \cap \mathcal{S})$ when $f$ is the indicator function of a subset $\mathcal{S}$ of $\mathbb{R}^2$.

Let us briefly describe the history of the problem. The Siegel transform was originally defined by Siegel [1945] as the sum over all nonzero lattice points for unimodular lattices of any rank. In the same paper Siegel proved a mean value theorem for the Siegel transform, which in the primitive set-up amounts to

$$\int_X \hat{f}(\Lambda)\, d\mu(\Lambda) = \frac{1}{\zeta(2)} \int_{\mathbb{R}^2} f(\boldsymbol{x})\, d\boldsymbol{x} \tag{1-10}$$

for any bounded compactly supported $f$ on $\mathbb{R}^2$. Since then there has been much work extending his result to higher moments. For example, Rogers [1955] proved a series of higher moment formulas, which in particular includes a second moment formula for the Siegel transform defined on the space of unimodular lattices of rank greater than 2. However, his result did not give a second moment formula on $X$ as in our setting. For this setting, Schmidt [1960] proved an upper bound for the second moment of the primitive Siegel transform of indicator functions on $\mathbb{R}^2$. His bound was later logarithmically improved by Randol [1970] for discs centered at the origin and by Athreya and Margulis [2009] for general indicator functions building on Randol's bound. Athreya and Konstantoulas [2016] obtained similar bounds on the space of general symplectic lattices for a certain family of indicator functions. Continuing [Athreya and Konstantoulas 2016], Kelmer and the second author [Kelmer and Yu 2019] proved a second moment formula on the space of symplectic lattices $Y_n := \mathrm{Sp}(2n, \mathbb{R})/\mathrm{Sp}(2n, \mathbb{Z})$. In particular, when $n = 1$ we have $Y_1 = X$ and their formula also applies to our setting.[1] However, for our applications all these formulas are not explicit enough.

We now state an explicit second moment formula which we use to derive Theorem 1.2.

**Theorem 1.5.** *For any $r \geq 0$ let $\mathcal{S}_r$ be the open square with vertices given by $(\pm e^{-r}, \pm e^{-r})$, and let $f_r$ be the indicator function of $\mathcal{S}_r$. Then we have*

$$\|\hat{f}_r\|_2^2 = \frac{8}{\zeta(2)} \left( e^{-2r} + \int_{\mathcal{D}_r} \left( \frac{e^{-r}}{x_1} + \frac{e^{-r}}{x_2} - \frac{1}{x_1 x_2} \right) dx_1\, dx_2 \right), \tag{1-11}$$

*where*

$$\mathcal{D}_r := \{\boldsymbol{x} = (x_1, x_2) \in \mathcal{S}_r \mid x_1 > 0,\ x_2 > 0,\ x_1 + x_2 > e^r\},$$

*and $\|\cdot\|_2$ stands for the $L^2$-norm with respect to $\mu$.*

[1] See also [Fairchild 2019] for moment formulas of the Siegel–Veech transform recently obtained by Fairchild.

**Remark 1.6.** When $r \geq (\log 2)/2$ the region $\mathcal{D}_r$ is empty, and (1-11) simply reads as

$$\|\hat{f}_r\|_2^2 = \frac{8e^{-2r}}{\zeta(2)}.$$

We note that the latter equality in fact already follows from Siegel's mean value theorem, since in this case for any unimodular lattice there can only be at most one pair of primitive lattice points allowed in $\mathcal{S}_r$, which implies that $\hat{f}_r/2$ is an indicator function on $X$. When $0 \leq r < (\log 2)/2$, the region $\mathcal{D}_r$ is not empty, and it is not hard to compute the integral in (1-11) explicitly; see (3-5) below. In particular, plugging $r = 0$ into (1-11) we have $\|\hat{f}_0\|_2^2 = (12/\pi)^2 - 8 \approx 6.59$.

In Section 2 we prove a much more general second moment formula, see Theorem 2.1, with an arbitrary bounded measurable subset $\mathcal{S}$ of $\mathbb{R}^2$ in place of $\mathcal{S}_r$. Theorem 1.5 is derived from Theorem 2.1 by taking $\mathcal{S} = \mathcal{S}_r$.

## 2. The second moment formula

In this section, we prove Theorem 1.5 by establishing the following second moment formula for quite general subsets of $\mathbb{R}^2$.

**Theorem 2.1.** *Let $\mathcal{S}$ be a measurable bounded subset of $\mathbb{R}^2$, and let $f$ be the indicator function of $\mathcal{S}$. Let $\widetilde{\mathcal{S}} = \{\boldsymbol{x} \in \mathbb{R}^2 \mid -\boldsymbol{x} \in \mathcal{S}\}$. Then we have*

$$\|\hat{f}\|_2^2 = \frac{1}{\zeta(2)}\left(\operatorname{area}(\mathcal{S}) + \operatorname{area}(\mathcal{S} \cap \widetilde{\mathcal{S}}) + \sum_{n \neq 0} \frac{\varphi(|n|)}{|n|} \int_{\mathcal{S}} |\mathcal{I}_{\boldsymbol{x}}^n| \, d\boldsymbol{x}\right),$$

*where $\varphi$ is the Euler's totient function, $\mathcal{I}_{\boldsymbol{x}}^n \subset \mathbb{R}$ is defined by*

$$\mathcal{I}_{\boldsymbol{x}}^n := \left\{ t \in \mathbb{R} \;\middle|\; n\left(\frac{-x_2}{x_1^2 + x_2^2}, \frac{x_1}{x_1^2 + x_2^2}\right) + t(x_1, x_2) \in \mathcal{S} \right\},$$

*and $|\mathcal{I}_{\boldsymbol{x}}^n|$ is the length of $\mathcal{I}_{\boldsymbol{x}}^n$ with respect to the Lebesgue measure on $\mathbb{R}$.*

Before giving the proof let us make a few remarks about Theorem 2.1. First we note that for any bounded $\mathcal{S}$ there exists a sufficiently large $T > 0$ depending on $\mathcal{S}$ such that for any $|n| > T$ the set $\mathcal{I}_{\boldsymbol{x}}^n$ is empty for all $\boldsymbol{x} \in \mathcal{S}$. Thus the series on the right-hand side of (2-1) is a finite sum. Next we note that if we further assume $\mathcal{S}$ is symmetric with respect to the origin, then by symmetry we have $\mathcal{S} \cap \widetilde{\mathcal{S}} = \mathcal{S}$ and $|\mathcal{I}_{\boldsymbol{x}}^n| = |\mathcal{I}_{\boldsymbol{x}}^{-n}|$ for any $n \neq 0$. In particular, for such $\mathcal{S}$ we have the slightly simpler formula

$$\|\hat{f}\|_2^2 = \frac{2}{\zeta(2)}\left(\operatorname{area}(\mathcal{S}) + \sum_{n=1}^{\infty} \frac{\varphi(n)}{n} \int_{\mathcal{S}} |\mathcal{I}_{\boldsymbol{x}}^n| \, d\boldsymbol{x}\right). \tag{2-1}$$

Finally we note that for any $\Lambda \in X$ and $f$ as in Theorem 2.1 we have

$$(\hat{f}(\Lambda))^2 = \hat{f}(\Lambda) + \hat{\chi}_{\mathcal{S} \cap \widetilde{\mathcal{S}}}(\Lambda) + \sum_{\substack{\boldsymbol{v}_1, \boldsymbol{v}_2 \in \Lambda_{\mathrm{pr}} \\ \text{lin. ind.}}} f(\boldsymbol{v}_1) f(\boldsymbol{v}_2).$$

Thus Theorem 2.1 together with (1-10) implies

$$\int_X \sum_{\substack{\boldsymbol{v}_1, \boldsymbol{v}_2 \in \Lambda_{\mathrm{pr}} \\ \text{lin. ind.}}} f(\boldsymbol{v}_1) f(\boldsymbol{v}_2) \, d\mu(\Lambda) = \frac{1}{\zeta(2)} \sum_{n \neq 0} \frac{\varphi(|n|)}{|n|} \int_{\mathcal{S}} |\mathcal{I}_{\boldsymbol{x}}^n| \, d\boldsymbol{x}. \tag{2-2}$$

It is worth pointing out that the above formula can be compared to its higher-dimensional analogue: when $f$ is an indicator function of a bounded measurable subset $\mathcal{S}$ of $\mathbb{R}^k$ with $k \geq 3$, $X = \mathrm{SL}_k(\mathbb{R})/\mathrm{SL}_k(\mathbb{Z})$, and $\mu$ is the Haar probability measure on $X$, according to Rogers' second moment formula [1955] the left-hand side of (2-2) equals $(\mathrm{vol}(\mathcal{S})/\zeta(k))^2$. However, as we can see here the $k = 2$ case is much more complicated, with the answer depending on both the shape and the position of $\mathcal{S}$.

***Coordinates and measures.*** We fix coordinates on $G = \mathrm{SL}_2(\mathbb{R})$ via the Iwasawa decomposition $G = KAN$ with

$$K = \{k_\theta \mid 0 \leq \theta < 2\pi\}, \quad A = \{a_s \mid s \in \mathbb{R}\}, \quad \text{and} \quad N = \{u_t \mid t \in \mathbb{R}\},$$

where

$$k_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \quad a_s = \begin{pmatrix} e^s & 0 \\ 0 & e^{-s} \end{pmatrix} \quad \text{and} \quad u_t = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$

Explicitly, under coordinates $g = k_\theta a_s u_t$, $\mu$ is given by

$$d\mu(g) = \frac{1}{\zeta(2)} e^{2s}\, d\theta\, ds\, dt. \tag{2-3}$$

There is a natural identification between the homogeneous space $G/N$ and $\mathbb{R}^2 \smallsetminus \{\mathbf{0}\}$ induced by the map $G \to \mathbb{R}^2 \smallsetminus \{\mathbf{0}\}$ sending $g = k_\theta a_s u_t \in G$ to

$$\boldsymbol{x}(g) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = g \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} e^s \cos\theta \\ e^s \sin\theta \end{pmatrix}, \tag{2-4}$$

the left column of $g$. The Lebesgue measure, $d\boldsymbol{x}$, on $\mathbb{R}^2 \smallsetminus \{\mathbf{0}\} \cong G/N$ can be expressed via the polar coordinates $(s, \theta)$ as

$$d\boldsymbol{x}(k_\theta a_s) = e^{2s}\, d\theta\, ds. \tag{2-5}$$

***The second moment formula.*** In this subsection we prove Theorem 2.1, and with some more analysis we prove Theorem 1.5. As the first step of our computation we recall the following preliminary identity which relies on a standard unfolding argument. We note that one can find it in [Lang 1975, Chapter VIII, Section 1], and we include a short proof here to make the paper self-contained. See also [Kelmer and Yu 2019, Proposition 2.3] for a generalization to the space of symplectic lattices.

**Lemma 2.2.** *For any bounded and compactly supported function $f$ on $\mathbb{R}^2$ and for any bounded $F \in L^2(X, \mu)$ we have*

$$\langle \hat{f}, F \rangle = \frac{1}{\zeta(2)} \int_{-\infty}^{\infty} \int_0^{2\pi} f(\boldsymbol{x}(k_\theta a_s)) \overline{\mathcal{P}_F(\boldsymbol{x}(k_\theta a_s))} e^{2s}\, d\theta\, ds,$$

*where $\mathcal{P}_F$ is defined by*

$$\mathcal{P}_F(\boldsymbol{x}(k_\theta a_s)) = \int_0^1 F(k_\theta a_s u_t \mathbb{Z}^2)\, dt$$

*with $k_\theta$, $a_s$ and $u_t$ as above, and $\langle \cdot, \cdot \rangle$ is the inner product on $L^2(X, \mu)$.*

*Proof.* Let $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ and let $\Gamma_\infty = \Gamma \cap N$. Recall that there is an identification between $\Gamma/\Gamma_\infty$ and $\mathbb{Z}^2_{\mathrm{pr}}$ sending $\gamma\Gamma_\infty$ to $\gamma\binom{1}{0}$. Using this identification, for any $\Lambda = g\mathbb{Z}^2$ with $g \in \mathrm{SL}_2(\mathbb{R})$ we can write

$$\hat{f}(\Lambda) = \sum_{\boldsymbol{v}\in\Lambda_{\mathrm{pr}}} f(\boldsymbol{v}) = \sum_{\boldsymbol{w}\in\mathbb{Z}^2_{\mathrm{pr}}} f(g\boldsymbol{w}) = \sum_{\gamma\in\Gamma/\Gamma_\infty} \tilde{f}(g\gamma), \tag{2-6}$$

where $\tilde{f}(g) := f\big(g\binom{1}{0}\big)$. We note that $\tilde{f}$ is a right $N$-invariant function on $G$. Let $\mathcal{F}_\Gamma$ be a fundamental domain for $X = G/\Gamma$, and let $\mathcal{F}_\infty$ be a fundamental domain for $G/\Gamma_\infty$. Note that using the Iwasawa decomposition $G = KAN$ we can choose

$$\mathcal{F}_\infty = \{k_\theta a_s u_t \mid 0 < \theta < 2\pi,\ s \in \mathbb{R},\ 0 < t < 1\}. \tag{2-7}$$

Moreover, fix a set of coset representatives $\Sigma_\infty \subset \Gamma$ for $\Gamma/\Gamma_\infty$, and note that $\bigcup_{\gamma\in\Sigma_\infty} \mathcal{F}_\Gamma\gamma$ is a disjoint union and forms a fundamental domain for $G/\Gamma_\infty$. Now for any bounded $F \in L^2(X,\mu)$, using (2-3), (2-6), (2-7) and the facts that $F$ is right $\Gamma$-invariant and $\tilde{f}$ is right $N$-invariant, we have

$$\begin{aligned}
\langle \hat{f}, F\rangle &:= \int_{\mathcal{F}_\Gamma} \hat{f}(g\mathbb{Z}^2)\overline{F(g\mathbb{Z}^2)}\,d\mu(g) = \sum_{\gamma\in\Gamma/\Gamma_\infty}\int_{\mathcal{F}_\Gamma} \tilde{f}(g\gamma)\overline{F(g\mathbb{Z}^2)}\,d\mu(g)\\
&= \sum_{\gamma\in\Sigma_\infty}\int_{\mathcal{F}_\Gamma\gamma} \tilde{f}(g)\overline{F(g\mathbb{Z}^2)}\,d\mu(g) = \int_{\bigsqcup_{\gamma\in\Sigma_\infty}\mathcal{F}_\Gamma\gamma} \tilde{f}(g)\overline{F(g\mathbb{Z}^2)}\,d\mu(g)\\
&= \int_{\mathcal{F}_\infty} \tilde{f}(g)\overline{F(g\mathbb{Z}^2)}\,d\mu(g) = \frac{1}{\zeta(2)}\int_{-\infty}^{\infty}\int_0^{2\pi}\int_0^1 \tilde{f}(k_\theta a_s u_t)\overline{F(k_\theta a_s u_t\mathbb{Z}^2)}e^{2s}\,dt\,d\theta\,ds\\
&= \frac{1}{\zeta(2)}\int_{-\infty}^{\infty}\int_0^{2\pi} f(\boldsymbol{x}(k_\theta a_s))\int_0^1 \overline{F(k_\theta a_s u_t\mathbb{Z}^2)}\,dt\,e^{2s}\,d\theta\,ds.
\end{aligned}$$

Finally, we note that the above equalities can be justified since $F$ is bounded and the defining series for $\hat{f}$ is absolutely convergent; see [Veech 1998, Lemma 16.10]. □

With this preliminary identity, we can now give:

*Proof of Theorem 2.1.* Using the relation (2-5) and Lemma 2.2 we have

$$\|\hat{f}\|_2^2 = \frac{1}{\zeta(2)}\int_{\mathbb{R}^2} f(\boldsymbol{x}(k_\theta a_s))\mathcal{P}_{\hat{f}}(\boldsymbol{x}(k_\theta a_s))\,d\boldsymbol{x} = \frac{1}{\zeta(2)}\int_{\mathcal{S}} \mathcal{P}_{\hat{f}}(\boldsymbol{x}(k_\theta a_s))\,d\boldsymbol{x}, \tag{2-8}$$

where

$$\mathcal{P}_{\hat{f}}(\boldsymbol{x}(k_\theta a_s)) = \int_0^1 \hat{f}(k_\theta a_s u_t\mathbb{Z}^2)\,dt,$$

with $k_\theta$, $a_s$ and $u_t$ as before. First, by the definition of the primitive Siegel transform we have

$$\hat{f}(k_\theta a_s u_t\mathbb{Z}^2) = \#\left\{(m,n) \in \mathbb{Z}^2_{\mathrm{pr}} \;\middle|\; k_\theta a_s u_t\binom{m}{n} \in \mathcal{S}\right\}.$$

Thus for $\boldsymbol{x}(k_\theta a_s) \in \mathcal{S}$ and $0 \le t < 1$ we have

$$\hat{f}(k_\theta a_s u_t\mathbb{Z}^2) = \sum_{(m,n)\in\mathbb{Z}^2_{\mathrm{pr}}} \chi_{I^{(m,n)}_{\boldsymbol{x}(k_\theta a_s)}}(t),$$

where

$$I^{(m,n)}_{\boldsymbol{x}(k_\theta a_s)} := \left\{ 0 \le t < 1 \;\middle|\; k_\theta a_s u_t \begin{pmatrix} m \\ n \end{pmatrix} \in \mathcal{S} \right\},$$

implying

$$\mathcal{P}_{\hat{f}}(\boldsymbol{x}(k_\theta a_s)) = \sum_{(m,n)\in\mathbb{Z}^2_{\mathrm{pr}}} |I^{(m,n)}_{\boldsymbol{x}(k_\theta a_s)}| = |I^{(1,0)}_{\boldsymbol{x}(k_\theta a_s)}| + |I^{(-1,0)}_{\boldsymbol{x}(k_\theta a_s)}| + \sum_{\substack{(m,n)\in\mathbb{Z}^2_{\mathrm{pr}} \\ n\neq 0}} |I^{(m,n)}_{\boldsymbol{x}(k_\theta a_s)}|.$$

Next, by direct computation we have for $\boldsymbol{x}(k_\theta a_s) = (x_1, x_2) = (e^s \cos\theta, e^s \sin\theta) \in \mathcal{S}$

$$k_\theta a_s u_t \begin{pmatrix} m \\ n \end{pmatrix} = n \begin{pmatrix} -e^{-s}\sin\theta \\ e^{-s}\cos\theta \end{pmatrix} + (m+nt) \begin{pmatrix} e^s\cos\theta \\ e^s \sin\theta \end{pmatrix} = n \begin{pmatrix} -x_2/(x_1^2+x_2^2) \\ x_1/(x_1^2+x_2^2) \end{pmatrix} + (m+nt) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \quad (2\text{-}9)$$

When $(m, n) = (1, 0)$ we have for $\boldsymbol{x}(k_\theta a_s) \in \mathcal{S}$

$$k_\theta a_s u_t \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

is contained in $\mathcal{S}$ for any $0 \le t < 1$. Thus $I^{(1,0)}_{\boldsymbol{x}(k_\theta a_s)} = [0, 1)$ and $|I^{(1,0)}_{\boldsymbol{x}(k_\theta a_s)}| = 1$ for any $\boldsymbol{x}(k_\theta a_s) \in \mathcal{S}$. Similarly, when $(m, n) = (-1, 0)$ we have for $\boldsymbol{x}(k_\theta a_s) \in \mathcal{S}$

$$k_\theta a_s u_t \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -x_1 \\ -x_2 \end{pmatrix}$$

is contained in $\mathcal{S}$ if and only if $\boldsymbol{x} \in \mathcal{S} \cap \widetilde{\mathcal{S}}$ with $\widetilde{\mathcal{S}}$ as in the theorem, implying $I^{(-1,0)}_{\boldsymbol{x}(k_\theta a_s)} = [0, 1)$ whenever $\boldsymbol{x} \in \mathcal{S} \cap \widetilde{\mathcal{S}}$.

When $n \neq 0$ by (2-9) we have for any integer $m$ coprime to $n$

$$\begin{aligned} |I^{(m,n)}_{\boldsymbol{x}}| &= \left| \left\{ 0 \le t < 1 \;\middle|\; n\left( \frac{-x_2}{x_1^2+x_2^2}, \frac{x_1}{x_1^2+x_2^2} \right) + (m+nt)(x_1, x_2) \in \mathcal{S} \right\} \right| \\ &= \left| \left\{ \frac{m}{n} \le t < 1 + \frac{m}{n} \;\middle|\; n\left( \frac{-x_2}{x_1^2+x_2^2}, \frac{x_1}{x_1^2+x_2^2} \right) + nt(x_1, x_2) \in \mathcal{S} \right\} \right|. \end{aligned}$$

We note that as $m$ runs through all the integers in each congruence class in $(\mathbb{Z}/|n|\mathbb{Z})^\times$, the intervals $[m/n, 1 + m/n)$ cover $\mathbb{R}$ exactly once. Thus for $n \neq 0$

$$\sum_{\substack{m\in\mathbb{Z} \\ (m,n)=1}} |I^{(m,n)}_{\boldsymbol{x}(k_\theta a_s)}| = \varphi(|n|) \left| \left\{ t \in \mathbb{R} \;\middle|\; n\left( \frac{-x_2}{x_1^2+x_2^2}, \frac{x_1}{x_1^2+x_2^2} \right) + nt(x_1, x_2) \in \mathcal{S} \right\} \right| = \frac{\varphi(|n|)}{|n|} |\mathcal{I}^n_{\boldsymbol{x}}|,$$

where $\varphi$ is the Euler's totient function and $\mathcal{I}^n_{\boldsymbol{x}}$ is as in Theorem 2.1. We thus have for $\boldsymbol{x} \in \mathcal{S}$

$$\mathcal{P}_{\hat{f}}(\boldsymbol{x}) = 1 + \chi_{\mathcal{S}\cap\widetilde{\mathcal{S}}}(\boldsymbol{x}) + \sum_{n\neq 0} \frac{\varphi(|n|)}{|n|} |\mathcal{I}^n_{\boldsymbol{x}}|.$$

We conclude the proof by plugging the above equation into (2-8). □

We can now give:

*Proof of Theorem 1.5.* To simplify notation for any $\boldsymbol{x} \in \mathbb{R}^2$, $t \in \mathbb{R}$, and $n \geq 1$ let

$$\boldsymbol{v}(\boldsymbol{x}, t, n) := n\left(\frac{-x_2}{x_1^2+x_2^2}, \frac{x_1}{x_1^2+x_2^2}\right) + t(x_1, x_2).$$

First we note that

$$\|\boldsymbol{v}(\boldsymbol{x}, t, n)\|_2^2 = \frac{n^2}{x_1^2+x_2^2} + t^2(x_1^2+x_2^2) \geq \frac{n^2}{x_1^2+x_2^2},$$

where $\|\cdot\|_2$ stands for the standard Euclidean norm on $\mathbb{R}^2$. Thus for $\boldsymbol{x} \in \mathcal{S}_r$ and $n \geq 2$ we have

$$\|\boldsymbol{v}(\boldsymbol{x}, t, n)\| \geq \frac{\sqrt{2}}{2}\|\boldsymbol{v}(\boldsymbol{x}, t, n)\|_2 \geq \frac{\sqrt{2}}{\|\boldsymbol{x}\|_2} > e^r \geq e^{-r},$$

implying that $\mathcal{I}_{\boldsymbol{x}}^n$ is empty for any $\boldsymbol{x} \in \mathcal{S}_r$ and any $n \geq 2$. Here $\|\cdot\|$ stands for the supremum norm on $\mathbb{R}^2$, and for the third inequality we used the fact that $\|\boldsymbol{x}\|_2 < \sqrt{2}e^{-r}$, which follows from $\boldsymbol{x}$ being an element of $\mathcal{S}_r$. Since $\mathcal{S}_r$ is symmetric with respect to the origin, applying (2-1) to $f = f_r$ we get

$$\|\hat{f}_r\|_2^2 = \frac{8e^{-2r}}{\zeta(2)} + \frac{2}{\zeta(2)}\int_{\mathcal{S}_r} |\mathcal{I}_{\boldsymbol{x}}^1|\, d\boldsymbol{x} = \frac{8e^{-2r}}{\zeta(2)} + \frac{8}{\zeta(2)}\int_{\mathcal{S}_r^+} |\mathcal{I}_{\boldsymbol{x}}^1|\, d\boldsymbol{x}, \tag{2-10}$$

where $\mathcal{S}_r^+$ is the intersection of $\mathcal{S}_r$ with the first quadrant, and for the second equality we used the fact that $|\mathcal{I}_{(x_1,x_2)}^1| = |\mathcal{I}_{(\pm x_1,\pm x_2)}^1|$ which follows from the invariance of $\mathcal{S}_r$ under reflections around the coordinate axes. We note that for $\boldsymbol{x} \in \mathcal{S}_r^+$

$$\left(\frac{-x_2}{x_1^2+x_2^2}, \frac{x_1}{x_1^2+x_2^2}\right) + t(x_1, x_2) \in \mathcal{S}_r$$

if and only if

$$-\frac{e^{-r}}{x_1} + \frac{x_2}{x_1(x_1^2+x_2^2)} < t < \frac{e^{-r}}{x_1} + \frac{x_2}{x_1(x_1^2+x_2^2)}$$

and

$$-\frac{e^{-r}}{x_2} - \frac{x_1}{x_2(x_1^2+x_2^2)} < t < \frac{e^{-r}}{x_2} - \frac{x_1}{x_2(x_1^2+x_2^2)}.$$

By direct computation if $r \geq (\log 2)/2$ then there is no $t \in \mathbb{R}$ satisfying above inequalities. Thus $\mathcal{I}_{\boldsymbol{x}}^1$ is empty, and the integral in the right-hand side of (2-10) is zero. If $0 \leq r < (\log 2)/2$, we define for any $\boldsymbol{x} \in \mathcal{S}_r^+$

$$L(\boldsymbol{x}) := \max\left\{-\frac{e^{-r}}{x_1} + \frac{x_2}{x_1(x_1^2+x_2^2)}, -\frac{e^{-r}}{x_2} - \frac{x_1}{x_2(x_1^2+x_2^2)}\right\},$$
$$U(\boldsymbol{x}) := \min\left\{\frac{e^{-r}}{x_1} + \frac{x_2}{x_1(x_1^2+x_2^2)}, \frac{e^{-r}}{x_2} - \frac{x_1}{x_2(x_1^2+x_2^2)}\right\}.$$

It is not hard to verify that as long as $0 \leq r < (\log 2)/2$, for $\boldsymbol{x} \in \mathcal{S}_r^+$ we have

$$L(\boldsymbol{x}) = -\frac{e^{-r}}{x_1} + \frac{x_2}{x_1(x_1^2+x_2^2)} \quad \text{and} \quad U(\boldsymbol{x}) = \frac{e^{-r}}{x_2} - \frac{x_1}{x_2(x_1^2+x_2^2)}.$$

Thus $\mathcal{I}_{\boldsymbol{x}}^1$ is nonempty if and only if $L(\boldsymbol{x}) < U(\boldsymbol{x})$ and whenever it is nonempty we have

$$\mathcal{I}_{\boldsymbol{x}}^1 = \left(-\frac{e^{-r}}{x_1} + \frac{x_2}{x_1(x_1^2+x_2^2)}, \frac{e^{-r}}{x_2} - \frac{x_1}{x_2(x_1^2+x_2^2)}\right).$$

By direct computation we have $L(\boldsymbol{x}) < U(\boldsymbol{x})$ if and only if $\boldsymbol{x} \in \mathcal{D}_r = \{(x_1, x_2) \in \mathcal{S}_r^+ \mid x_1 + x_2 > e^r\}$. Hence

$$\begin{aligned}\|\hat{f}_r\|_2^2 &= \frac{8e^{-2r}}{\zeta(2)} + \frac{8}{\zeta(2)} \int_{\mathcal{D}_r} \left(\left(\frac{e^{-r}}{x_2} - \frac{x_1}{x_2(x_1^2+x_2^2)}\right) - \left(-\frac{e^{-r}}{x_1} + \frac{x_2}{x_1(x_1^2+x_2^2)}\right)\right) dx_1\, dx_2 \\ &= \frac{8e^{-2r}}{\zeta(2)} + \frac{8}{\zeta(2)} \int_{\mathcal{D}_r} \left(\frac{e^{-r}}{x_1} + \frac{e^{-r}}{x_2} - \frac{1}{x_1 x_2}\right) dx_1\, dx_2. \end{aligned}$$

□

Besides the sets $\mathcal{S}_r$, another natural candidate to test formula (2-1) is the family of indicator functions of balls. For any $R > 0$ let $\mathcal{B}_R$ be the open ball of radius $R$ centered at the origin, and let $h_R$ be the indicator function of $\mathcal{B}_R$. We note that [Randol 1970] established an asymptotic formula for $\|\hat{h}_R\|_2^2$ for large $R$, and here we prove the following formula for $\|\hat{h}_R\|_2^2$:

**Corollary 2.3.** *For any $R > 0$ let $h_R$ be as above. Then we have*

$$\|\hat{h}_R\|_2^2 = \frac{12R^2}{\pi} + \frac{48}{\pi} \sum_{n=1}^{\lfloor R^2 \rfloor} \varphi(n) \left(\frac{\sqrt{R^4 - n^2}}{n} + \arcsin\left(\frac{n}{R^2}\right) - \frac{\pi}{2}\right).$$

*Proof.* Since $\mathcal{B}_R$ is symmetric with respect to the origin, we can apply (2-1) to $\|\hat{h}_R\|_2^2$, and use $\zeta(2) = \pi^2/6$ to get

$$\|\hat{h}_R\|_2^2 = \frac{12R^2}{\pi} + \frac{12}{\pi^2} \sum_{n=1}^{\infty} \frac{\varphi(n)}{n} \int_{\mathcal{B}_R} |\mathcal{I}_{\boldsymbol{x}}^n|\, d\boldsymbol{x},$$

where

$$\mathcal{I}_{\boldsymbol{x}}^n := \left\{ t \in \mathbb{R} \;\middle|\; \left\| n\left(\frac{-x_2}{x_1^2+x_2^2}, \frac{x_1}{x_1^2+x_2^2}\right) + t(x_1, x_2) \right\|_2 < R \right\}.$$

Using the polar coordinates, for any $(x_2, x_2) = (r\cos\theta, r\sin\theta) \in \mathcal{B}_R$ and $n \geq Rr$ we can write

$$\left\| n\left(\frac{-x_2}{x_1^2+x_2^2}, \frac{x_1}{x_1^2+x_2^2}\right) + t(x_1, x_2) \right\|_2^2 = \frac{n^2}{r^2} + t^2 r^2 \geq R^2,$$

implying that $\mathcal{I}_{\boldsymbol{x}}^n$ is empty whenever $n \geq Rr = R\|\boldsymbol{x}\|_2$. In particular, $\mathcal{I}_{\boldsymbol{x}}^n$ is empty for any $\boldsymbol{x} \in \mathcal{B}_R$ if $n \geq R^2$. Similarly, for any $1 \leq n \leq \lfloor R^2 \rfloor$ the set $\mathcal{I}_{\boldsymbol{x}}^n$ is empty if $\|\boldsymbol{x}\|_2 \leq n/R$, and

$$\mathcal{I}_{\boldsymbol{x}}^n = \left(-\frac{\sqrt{R^2r^2 - n^2}}{r^2}, \frac{\sqrt{R^2r^2 - n^2}}{r^2}\right)$$

if $n/R < \|\boldsymbol{x}\|_2 < R$. Hence

$$\|\hat{h}_R\|_2^2 = \frac{12R^2}{\pi} + \frac{12}{\pi^2} \sum_{n=1}^{\lfloor R^2 \rfloor} \frac{\varphi(n)}{n} \int_0^{2\pi} \int_{n/R}^{R} \frac{2\sqrt{R^2r^2 - n^2}}{r^2} r\, dr\, d\theta$$

$$= \frac{12R^2}{\pi} + \frac{48}{\pi}\sum_{n=1}^{\lfloor R^2\rfloor}\varphi(n)\int_1^{R^2/n}\sqrt{1-r^{-2}}\,dr$$
$$= \frac{12R^2}{\pi} + \frac{48}{\pi}\sum_{n=1}^{\lfloor R^2\rfloor}\varphi(n)\left(\frac{\sqrt{R^4-n^2}}{n} + \arcsin\left(\frac{n}{R^2}\right) - \frac{\pi}{2}\right),$$

where for the second equality we applied a change of variable $(R/n)r \mapsto r$, and for the last equality we used the fact that $\int \sqrt{1-r^{-2}}\,dr = \sqrt{r^2-1} + \arcsin(1/r) + C$ for $r \geq 1$. □

## 3. Measure estimates of the shrinking targets

In this section, using the methods developed in the previous section, we prove Theorem 1.2 and then use it to derive Theorem 1.3 and Corollary 1.4.

*Proof of Theorem 1.2.* For any $r > 0$, let $f_r$ be the indicator function of $\mathcal{S}_r$ as before. For any integer $k \geq 0$, let $B_r^k \subset X$ be the set of unimodular lattices having $2k$ nonzero primitive points in $\mathcal{S}_r$. First, we note that $K_r = B_r^0$ consists of lattices with no nonzero points in $\mathcal{S}_r$. Moreover, for any $\Lambda \in X$, there are at most two linearly independent primitive points of $\Lambda$ inside $\mathcal{S}_r$. We thus have for any $r > 0$

$$\sum_{k=0}^{2}\mu(B_r^k) = 1, \tag{3-1}$$

and

$$\hat{f}_r = 2\chi_{B_r^1} + 4\chi_{B_r^2}.$$

Thus we can take the first moment and apply (1-10) to get

$$\mu(B_r^1) + 2\mu(B_r^2) = \frac{1}{2}\int_X \hat{f}_r(\Lambda)\,d\mu(\Lambda) = \frac{2e^{-2r}}{\zeta(2)}. \tag{3-2}$$

Taking the second moment of $\hat{f}_r$ we get

$$4\mu(B_r^1) + 16\mu(B_r^2) = \|\hat{f}_r\|_2^2. \tag{3-3}$$

Solving (3-1), (3-2) and (3-3) and applying Theorem 1.5 to (3-3), we get

$$\mu(K_r) = \mu(B_r^0) = 1 - \frac{2e^{-2r}}{\zeta(2)} + \frac{1}{\zeta(2)}\int_{\mathcal{D}_r}\left(\frac{e^{-r}}{x_1} + \frac{e^{-r}}{x_2} - \frac{1}{x_1x_2}\right)dx_1\,dx_2.$$

By direct computation we have for $0 < r < \frac{1}{2}\log 2$

$$\int_{\mathcal{D}_r}\left(\frac{e^{-r}}{x_1} + \frac{e^{-r}}{x_2} - \frac{1}{x_1x_2}\right)dx_1\,dx_2$$
$$= 2(1-r)(2e^{-2r}-1+r) + (2-2e^{-2r}-2r)\log(1-e^{-2r}) - 2r^2 + \int_{1-e^{-2r}}^{e^{-2r}}\frac{\log t}{1-t}\,dt$$
$$= 2(1-r)(2e^{-2r}-1+r) + (2-2e^{-2r}-2r)\log(1-e^{-2r}) - 2r^2 + \mathrm{Li}_2(1-e^{-2r}) - \mathrm{Li}_2(e^{-2r}), \tag{3-4}$$

where $\mathrm{Li}_s(z) = \sum_{k=1}^{\infty} z^k/k^s$ is the polylogarithm function. Now for the term $\log(1-e^{-2r})$, using the Taylor expansion $e^{-2r} = 1 - 2r + 2r^2 + O(r^3)$, we get

$$\log(1-e^{-2r}) = \log(2r) + \log(1 - r + O(r^2)) = \log(2r) - r + O(r^2).$$

Using the series representation $\mathrm{Li}_2(z) = \sum_{k=1}^{\infty} z^k/k^2$, we get $\mathrm{Li}_2(1-e^{-2r}) = 2r + O(r^2)$. Finally for the term $\mathrm{Li}_2(e^{-2r})$ we have the expansion, see [Wood 1992, Equation (9.7)],

$$\mathrm{Li}_2(e^{-2r}) = -2r(1 - \log(2r)) + \zeta(2) + O(r^2).$$

Plugging these into (3-4) and using the expansion $e^{-2r} = 1 - 2r + 2r^2 + O(r^3)$, we get

$$\int_{\mathcal{D}_r} \left( \frac{e^{-r}}{x_1} + \frac{e^{-r}}{x_2} - \frac{1}{x_1 x_2} \right) dx_1\, dx_2 = 2 - \zeta(2) - 4r - 4r^2 \log r + O(r^2), \tag{3-5}$$

implying

$$\mu(K_r) = 1 - \frac{2e^{-2r}}{\zeta(2)} + \frac{1}{\zeta(2)}(2 - \zeta(2) - 4r - 4r^2 \log r + O(r^2)) = -\frac{4r^2 \log r}{\zeta(2)} + O(r^2),$$

finishing the proof. □

To estimate the measure of the thickening, we will need the following two preliminary lemmas. We note that by the Hajós–Minkowski theorem, see [Cassels 1997, IX.1.3], we have

$$K_0 = \Delta^{-1}\{0\} = \bigcup_{x\in[0,1)} \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \mathbb{Z}^2 \bigcup \begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix} \mathbb{Z}^2.$$

A simple observation is that any $\Lambda \in K_0$ contains either the point $(1, 0)$ or the point $(0, 1)$. Thus intuitively one shall expect that when $r$ is small, lattices in $K_r$ contain points close to either $(1, 0)$ or $(0, 1)$. For any $r > 0$, let $\mathcal{A}_r \subset \mathbb{R}^2$ be the closed rectangle with vertices $(\pm\sqrt{e^{2r}-1}, e^r)$ and $(\pm\sqrt{e^{2r}-1}, e^{-r})$ and let $\mathcal{C}_r$ be the closed rectangle with vertices $(e^r, \pm\sqrt{e^{2r}-1})$ and $(e^{-r}, \pm\sqrt{e^{2r}-1})$; see Figure 1. The following lemma asserts that when $r$ is small, then any $\Lambda \in K_r$ contains points either in $\mathcal{A}_r$ or in $\mathcal{C}_r$ (noting that $\mathcal{A}_r$ is a small rectangle containing $(0, 1)$ and $\mathcal{C}_r$ is a small rectangle containing $(1, 0)$).

**Lemma 3.1.** *Let $\mathcal{A}_r$ and $\mathcal{C}_r$ be as above. For any $0 < r < \log 1.01$ and for any $\Lambda \in K_r$, we have $\Lambda_{\mathrm{pr}} \cap (\mathcal{A}_r \cup \mathcal{C}_r) \neq \varnothing$.*

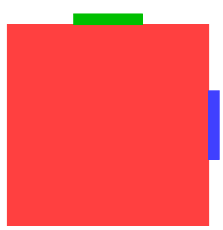


**Figure 1.** The square $\mathcal{S}_r$ (red), the rectangles $\mathcal{A}_r$ (green) and $\mathcal{C}_r$ (blue).

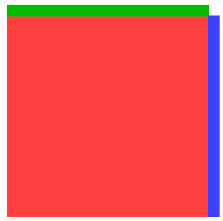


**Figure 2.** The square $\mathcal{S}_r$ (red), the rectangles $\mathcal{U}_r$ (green) and $\mathcal{R}_r$ (blue).

*Proof.* Let $\mathcal{U}_r$ be the closed rectangle with vertices $(\pm e^{-r}, e^{-r})$ and $(\pm e^{-r}, e^{r})$, and let $\mathcal{R}_r$ be the closed rectangle with vertices $(e^{-r}, \pm e^{-r})$ and $(e^{r}, \pm e^{-r})$; see Figure 2. Let

$$\widetilde{\mathcal{U}}_r := \{\boldsymbol{x} \in \mathbb{R}^2 \mid -\boldsymbol{x} \in \mathcal{U}_r\}.$$

Consider the rectangle $\mathcal{U}_r \sqcup \mathcal{S}_r \sqcup \widetilde{\mathcal{U}}_r$ and note that it has area 4. For any $\varepsilon > 0$ let $\mathcal{U}_{r,\varepsilon}$ be the open rectangle with vertices $(\pm e^{-r}, \pm(e^{r} + \varepsilon))$. Applying the Minkowski's convex body theorem to $\mathcal{U}_{r,\varepsilon}$ and letting $\varepsilon$ approach zero, we see that for any $\Lambda \in X$, $\Lambda_{\mathrm{pr}}$ intersects $\mathcal{U}_r \sqcup \mathcal{S}_r \sqcup \widetilde{\mathcal{U}}_r$ nontrivially. Now let $\Lambda \in K_r$; since $\Lambda$ has no nonzero point in $\mathcal{S}_r$ and $\Lambda_{\mathrm{pr}}$ is invariant under inversion, we have $\Lambda_{\mathrm{pr}} \cap \mathcal{U}_r \neq \varnothing$. Similarly we also have $\Lambda_{\mathrm{pr}} \cap \mathcal{R}_r \neq \varnothing$. Moreover, we note that for $0 < r < \log 1.01$, we have $\Lambda \cap \mathcal{U}_r = \Lambda_{\mathrm{pr}} \cap \mathcal{U}_r$ and $\Lambda \cap \mathcal{R}_r = \Lambda_{\mathrm{pr}} \cap \mathcal{R}_r$. This is because otherwise there would be some nonzero point $\boldsymbol{v} \in \Lambda \cap (\mathcal{U}_r \cup \mathcal{R}_r)$ and some integer $k \geq 2$ such that $\boldsymbol{v}/k \in \Lambda_{\mathrm{pr}}$, but $\boldsymbol{v} \in \mathcal{U}_r \cup \mathcal{R}_r$ and $k \geq 2$ imply that $\boldsymbol{v}/k \in \mathcal{S}_r$, contradicting the assumption that $\Lambda_{\mathrm{pr}} \cap \mathcal{S}_r = \varnothing$. Let $\boldsymbol{v}_1 = (t_1, 1 + v_1)$ be a point in $\Lambda_{\mathrm{pr}} \cap \mathcal{U}_r$ that is closest to the $y$-axis and let $\boldsymbol{v}_2 = (1 + v_2, t_2)$ be a point in $\Lambda_{\mathrm{pr}} \cap \mathcal{R}_r$ that is closest to the $x$-axis. We thus have $|t_i| \leq e^{-r}$ and $e^{-r} \leq 1 + v_i \leq e^{r}$ for $i = 1, 2$.

Let $\mathcal{P}_{\boldsymbol{v}_1, \boldsymbol{v}_2}$ be the parallelogram spanned by $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$. Then we have for $0 < r < \log 1.01$

$$|\mathcal{P}_{\boldsymbol{v}_1, \boldsymbol{v}_2}| = |(1 + v_1)(1 + v_2) - t_1 t_2| = (1 + v_1)(1 + v_2) - t_1 t_2 \leq e^{2r} + e^{-2r} < 3,$$

where $|\mathcal{P}_{\boldsymbol{v}_1, \boldsymbol{v}_2}|$ denotes the area of $\mathcal{P}_{\boldsymbol{v}_1, \boldsymbol{v}_2}$, and for the second equality we used that

$$(1 + v_1)(1 + v_2) \geq e^{-2r} \geq |t_1 t_2|.$$

Thus $|\mathcal{P}_{\boldsymbol{v}_1, \boldsymbol{v}_2}|$ equals 1 or 2. We claim that $|\mathcal{P}_{\boldsymbol{v}_1, \boldsymbol{v}_2}| = 1$. Suppose not; then $|\mathcal{P}_{\boldsymbol{v}_1, \boldsymbol{v}_2}| = 2$ and we have for $0 < r < \log 1.01$

$$t_1 t_2 = v_1 + v_2 + v_1 v_2 - 1 \leq 2(e^{r} - 1) + (e^{r} - 1)^2 - 1 < 0$$

and

$$|t_1 t_2| = 1 - v_1 - v_2 - v_1 v_2 \geq 1 - 2(e^{r} - 1) - (e^{r} - 1)^2 = 2 - e^{2r} > 0.9.$$

This implies $\min\{|t_1|, |t_2|\} > 0.9/e^{-r} > 0.9$. Since $t_1 t_2 < 0$, without loss of generality we may assume that $t_2 < 0$. Then we have $-e^{-r} \leq t_2 < -0.9$. On one hand, since $|\mathcal{P}_{\boldsymbol{v}_1, \boldsymbol{v}_2}| = 2$ and $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \Lambda_{\mathrm{pr}}$, we have

$$\boldsymbol{w} := \frac{\boldsymbol{v}_1 + \boldsymbol{v}_2}{2} = \left( \frac{t_1 + 1 + v_2}{2}, \frac{t_2 + 1 + v_1}{2} \right) \in \Lambda.$$

On the other hand, we have

$$0 < \frac{t_1 + 1 + v_2}{2} \le \frac{e^{-r} + e^r}{2} < e^r, \quad 0 < \frac{t_2 + 1 + v_1}{2} < \frac{1 + v_1}{2} \le \frac{e^r}{2} < e^{-r},$$

and $\boldsymbol{w} \notin \mathcal{S}_r$ implying $\boldsymbol{w} \in \mathcal{R}_r$. Thus $\boldsymbol{w} \in \Lambda \cap \mathcal{R}_r = \Lambda_{\mathrm{pr}} \cap \mathcal{R}_r$ is also a primitive vector of $\Lambda$. Moreover, since $-e^{-r} \le t_2 < -0.9$, we have

$$0 < \frac{t_2 + 1 + v_1}{2} < \frac{e^r - 0.9}{2} < \frac{1.01 - 0.9}{2} = 0.055 < |t_2|,$$

contradicting the assumption that $\boldsymbol{v}_2$ is the closest point in $\Lambda_{\mathrm{pr}} \cap \mathcal{R}_r$ to the $x$-axis. We thus have proved the claim, and it implies

$$|t_1 t_2| = |v_1 + v_2 + v_1 v_2| \le 2(e^r - 1) + (e^r - 1)^2 = e^{2r} - 1.$$

Hence we have

$$\min\{|t_1|, |t_2|\} \le \sqrt{|t_1 t_2|} \le \sqrt{e^{2r} - 1},$$

which implies $\Lambda_{\mathrm{pr}} \cap (\mathcal{A}_r \cup \mathcal{C}_r) \neq \varnothing$ finishing the proof. □

The following lemma states that for $r > 0$ small, the orbits $a_s K_r$ will completely leave the set $K_r$ very shortly, and will remain separated for quite a long time.

**Lemma 3.2.** *For any* $0 < r < \log 1.01$ *and any* $6r \le |s| \le \log 1.9$, *we have*

$$a_s K_r \cap K_r = \varnothing.$$

*Proof.* Suppose not, then there exists some $\Lambda \in a_s K_r \cap K_r$, and by definition the intersection of $\Lambda_{\mathrm{pr}}$ with $\mathcal{S}_r \cup a_s \mathcal{S}_r$ is empty. Without loss of generality we may assume that $s > 0$. By Lemma 3.1 we have $\Lambda_{\mathrm{pr}} \cap (\mathcal{A}_r \cup \mathcal{C}_r) \neq \varnothing$ and similarly, $\Lambda_{\mathrm{pr}} \cap (a_s \mathcal{A}_r \cup a_s \mathcal{C}_r) \neq \varnothing$. We note that $a_s \mathcal{A}_r$ is the rectangle with vertices $(\pm e^s \sqrt{e^{2r} - 1}, e^{r-s})$ and $(\pm e^s \sqrt{e^{2r} - 1}, e^{-r-s})$. Since $e^{6r} \le e^s \le 1.9$ we have $a_s \mathcal{A}_r \subseteq \mathcal{S}_r$ implying $\Lambda_{\mathrm{pr}} \cap a_s \mathcal{C}_r \neq \varnothing$. Similarly, we have $\mathcal{C}_r \subseteq a_s \mathcal{S}_r$ and this implies $\Lambda_{\mathrm{pr}} \cap \mathcal{A}_r \neq \varnothing$ (see Figure 4). Let $\boldsymbol{v}_1 \in \Lambda_{\mathrm{pr}} \cap \mathcal{A}_r$ and $\boldsymbol{v}_2 \in \Lambda_{\mathrm{pr}} \cap a_s \mathcal{C}_r$, and let $\mathcal{P}_{\boldsymbol{v}_1, \boldsymbol{v}_2}$ be the parallelogram spanned by $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$. Then for $0 < r < \log 1.01$ and $6r \le s \le \log 1.9$ we have

$$1 < e^{s-2r} - (e^{2r} - 1)e^{-s} \le |\mathcal{P}_{\boldsymbol{v}_1, \boldsymbol{v}_2}| \le e^{s+2r} + (e^{2r} - 1)e^{-s} < 2$$

contradicting the fact that $|\mathcal{P}_{\boldsymbol{v}_1, \boldsymbol{v}_2}|$ is a positive integer. □

We can now give:

*Proof of Theorem 1.3.* We prove the upper and lower bounds separately. For the upper bound, we first note that for any $\boldsymbol{v} \in \mathbb{R}^2$ we have $e^{-|s|} \|\boldsymbol{v}\| \le \|a_s \boldsymbol{v}\| \le e^{|s|} \|\boldsymbol{v}\|$. Hence for any $\Lambda \in X$ we have

$$|\Delta(a_s \Lambda) - \Delta(\Lambda)| \le |s|.$$

This implies that for any $s \in \mathbb{R}$ and any $r > 0$

$$a_s K_r \subset K_{r+|s|}. \tag{3-6}$$

**Figure 3.** Figure 1 under the flow $a_s$: the rectangles $a_s\mathcal{S}_r$ (orange), $a_s\mathcal{A}_r$ (brown), and $a_s\mathcal{C}_r$ (purple).

**Figure 4.** Figures 1 and 3 in one picture: the rectangle $a_s\mathcal{A}_r$ (brown) is contained in $\mathcal{S}_r$ (red), the rectangle $\mathcal{C}_r$ (blue) is contained in $a_s\mathcal{S}_r$ (orange).

Let $N = \lceil 1/r \rceil$. Using (3-6) and the fact that $1/N \le r$ we can estimate

$$\bigcup_{0\le s<1} a_{-s}K_r = \bigcup_{0\le i<N}\ \bigcup_{0\le t<1/N} a_{-i/N}a_{-t}K_r \subset \bigcup_{0\le i<N} a_{-i/N}K_{2r}.$$

Hence by Theorem 1.2 and since $N \asymp 1/r$ we have

$$\mu\Big(\bigcup_{0\le s<1} a_{-s}K_r\Big) \le \sum_{i=0}^{N-1} \mu(a_{-i/N}K_{2r}) \asymp r\log\Big(\frac{1}{r}\Big).$$

For the lower bound, for $0 < r < \log 1.01$ let $N = \lfloor 1/(6r) \rfloor$. First we have

$$\bigcup_{0\le i<\lfloor N\log 1.9\rfloor} a_{-i/N}K_r \subseteq \bigcup_{0\le s<1} a_{-s}K_r.$$

Moreover, for each $0 \le i < j < \lfloor N\log 1.9\rfloor$, we have $6r \le 1/N \le (j-i)/N < \log 1.9$; thus by Lemma 3.2 we have

$$a_{-i/N}K_r \cap a_{-j/N}K_r = a_{-j/N}(a_{(j-i)/N}K_r \cap K_r) = \varnothing.$$

Thus the union $\bigcup_{0\le i<\lfloor N\log 1.9\rfloor} a_{-i/N}K_r$ is disjoint and, again applying Theorem 1.2 and noting that $N \asymp 1/r$ we can estimate

$$\mu\Big(\bigcup_{0\le s<1} a_{-s}K_r\Big) \ge \sum_{i=0}^{\lfloor N\log 1.9\rfloor-1} \mu(a_{-i/N}K_r) \asymp r\log\Big(\frac{1}{r}\Big),$$

finishing the proof. □

*Proof of Corollary 1.4.* First we note that we can assume $\lim_{s\to\infty} r(s) = 0$ since otherwise both series would diverge. It follows that there exists $N > 0$ such that for any $n > N$, $0 < r(n) < \log 1.01$. Next, since $r(\cdot)$ is nonincreasing, for any $n > N$ we have

$$\bigcup_{0\le s<1} a_{-s} K_{r(n+1)} \subset \widetilde{B}_n \subset \bigcup_{0\le s<1} a_{-s} K_{r(n)}.$$

Moreover, since $n > N$ we have $0 < r(n+1) \le r(n) < \log 1.01$. Applying Theorem 1.3 to the left- and right-hand sides of the above inclusion relations we get

$$r(n+1)\log\left(\frac{1}{r(n+1)}\right) \ll \mu(\widetilde{B}_n) \ll r(n)\log\left(\frac{1}{r(n)}\right),$$

which finishes the proof. □

## 4. The dynamical Borel–Cantelli lemma

In this section we give the proof of Theorem 1.1 based on Theorem KW. Recall that for a given function $\psi : [t_0, \infty) \to (0, \infty)$ with $t_0 \ge 1$ fixed, we say a real number $x \in \mathbb{R}$ is *$\psi$-Dirichlet* if the system of inequalities

$$|qx - p| < \psi(t) \quad \text{and} \quad |q| < t$$

has a solution in $(p, q) \in \mathbb{Z} \times (\mathbb{Z} \smallsetminus \{0\})$ for all sufficiently large $t$. Let us denote by $D(\psi)$ the set of all $\psi$-Dirichlet numbers. Theorem KW gives a zero-one law for the Lebesgue measure of $D(\psi)$ as follows: if $\psi : [t_0, \infty) \to (0, \infty)$ is a continuous, nonincreasing function satisfying (1-6) and (1-7), then the series (1-8) diverges (resp. converges) if and only if the Lebesgue measure of $D(\psi)$ (resp. of $D(\psi)^c$) is zero.

For our purpose, we prove the following slightly modified version of *Dani correspondence*.

**Lemma 4.1.** *Let $\psi : [t_0, \infty) \to (0, \infty)$ be a continuous, nonincreasing function satisfying (1-6) and (1-7). Then there exists a unique continuous, nonincreasing function*

$$r = r_\psi : [s_0, \infty) \to (0, \infty), \quad \text{where } s_0 = \frac{\log t_0}{2} - \frac{\log \psi(t_0)}{2}$$

*such that*

$$\text{the function } s \mapsto s + r(s) \text{ is nondecreasing}, \tag{4-1}$$

*and*

$$\psi(e^{s-r(s)}) = e^{-s-r(s)} \quad \text{for all } s \ge s_0. \tag{4-2}$$

*Conversely, given a continuous, nonincreasing function $r : [s_0, \infty) \to (0, \infty)$ satisfying (4-1), then there exists a unique continuous, nonincreasing function $\psi = \psi_r : [t_0, \infty) \to (0, \infty)$ with $t_0 = e^{s_0 - r(s_0)}$ satisfying (1-6), (1-7) and (4-2). Furthermore, if we assume $\lim_{t\to\infty} t\psi(t) = 1$ (or equivalently, $\lim_{s\to\infty} r(s) = 0$), then the series in (1-8) diverges if and only if the series*

$$\sum_n r(n)\log\left(\frac{1}{r(n)}\right) \tag{4-3}$$

*diverges.*

*Proof.* The correspondence between $\psi = \psi_r$ and $r = r_\psi$ follows from the exact same construction as in [Kleinbock and Margulis 1999, Lemma 8.3], where $\psi(\cdot)$ and $r(\cdot)$ determine each other with

the relations

$$e^s\psi(t) = e^{-r(s)} = e^{-s}t,$$

with $s$ and $t$ satisfying $s = (\log t)/2 - (\log \psi(t))/2$. The only difference is that here we require the two extra assumptions (1-6) and (1-7) on $\psi$ which are respectively equivalent to the assumptions that $r(\cdot)$ is nonincreasing and $r(\cdot)$ is positive. We refer the reader to [Kleinbock and Margulis 1999, Lemma 8.3] for more details about this correspondence.

For the furthermore part, first we claim that the series in (1-8) diverges if and only if the integral

$$\int_{t_0}^{\infty} \frac{-(1-t\psi(t))\log(1-t\psi(t))}{t}\,dt \tag{4-4}$$

diverges. It suffices to show the function $G(t) := -\log(1-t\psi(t))(1-t\psi(t))$ is eventually nonincreasing in $t$. Note that the function $T \mapsto -T\log T$ is strictly increasing on the interval $(0, e^{-1})$. Since $\lim_{t\to\infty} t\psi(t) = 1$ and $t\psi(t) < 1$ for all $t \geq t_0$, there exists some $T_0 > t_0$ such that for all $t > T_0$, $0 < 1 - t\psi(t) < e^{-1}$. Moreover, together with the assumption (1-6) we get that $G(t)$ is nonincreasing in $t$ for any $T > T_0$, finishing the proof the claim. Next, since $r(\cdot)$ is positive and nonincreasing, we have $0 < r(s) \leq r(s_0)$. Thus there exist constants $0 < c_1 < c_2$ such that for all $s \geq s_0$ and all $t \geq t_0$ with $s = (\log t)/2 - (\log \psi(t))/2$ we have

$$c_1 r(s) \leq 1 - t\psi(t) = 1 - e^{-2r(s)} \leq c_2 r(s).$$

This also implies

$$-\log(1-t\psi(t)) = -\log(r(s)) + O_{c_1,c_2}(1) \asymp_{c_1,c_2} -\log(r(s)),$$

where for the second estimate we used that $\lim_{s\to\infty} r(s) = 0$. Moreover, since $r(\cdot)$ is nonincreasing and continuous, it is differentiable at Lebesgue almost every $s \in \mathbb{R}$, and we denote by $r'(s)$ its derivative at $s \in \mathbb{R}$ whenever it exists. Using the relation $t = e^{s-r(s)}$ we get $dt/t = (1 - r'(s))\,ds$ for Lebesgue almost every $s \in \mathbb{R}$. We thus have

$$\int_{t_0}^{\infty} \frac{-(1-t\psi(t))\log(1-t\psi(t))}{t}\,dt \asymp_{c_1,c_2} \int_{s_0}^{\infty} -r(s)\log(r(s))(1-r'(s))\,ds \asymp \int_{s_0}^{\infty} -r(s)\log(r(s))\,ds,$$

where for the second estimate we used that $1 \leq 1 - r'(s) \leq 2$ for Lebesgue almost every $s \in \mathbb{R}$ which comes from the assumption (4-1) and that $r(\cdot)$ is nonincreasing. Finally, we conclude the proof by noting that the integral $\int_{s_0}^{\infty} -r(s)\log(r(s))\,ds$ diverges if and only if the series $\sum_n -r(n)\log(r(n))$ diverges since $\lim_{s\to\infty} r(s) = 0$ and $r(\cdot)$ is nonincreasing, which imply that the function $s \mapsto -r(s)\log(r(s))$ is eventually nonincreasing in $s$. □

As mentioned in the Introduction, we have the following dynamical interpretation of $\psi$-Dirichlet numbers.

**Lemma 4.2** [Kleinbock and Wadleigh 2018, Proposition 4.5]. *Let $\psi : [t_0, \infty) \to (0, \infty)$ be a continuous and nonincreasing function satisfying (1-6) and (1-7). Let $r = r_\psi$ be as in Lemma 4.1. Then $x \in D(\psi)^c$ if and only if*

$$a_s \Lambda_x \in K_{r(s)} \quad \textit{for an unbounded set of } s, \tag{4-5}$$

*where* $a_s = \operatorname{diag}(e^s, e^{-s})$ *and*

$$\Lambda_x = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \mathbb{Z}^2 \in X$$

*are as before.*

Combining Theorem KW with Lemmas 4.1 and 4.2, we immediately have the following zero-one law.

**Proposition 4.3.** *Let* $r : [s_0, \infty) \to (0, \infty)$ *be continuous, nonincreasing, satisfying (4-1) and such that* $\lim_{s\to\infty} r(s) = 0$. *Then (4-5) holds for Lebesgue almost every (resp. almost no)* $x \in \mathbb{R}$ *provided that the series (4-3) diverges (resp. converges).*

To connect the above proposition with the corresponding property of almost every $\Lambda \in X$, we need an auxiliary lemma, which borrows some ideas from the work [Kleinbock and Rao 2019] of the first author with Anurag Rao.

**Lemma 4.4.** *Let* $r(\cdot)$ *be as in Proposition 4.3. For any* $c \in \mathbb{R}$ *and* $\lambda > 0$ *let*

$$r_{c,\lambda}(s) := r(s + c) - \lambda e^{-2(s+c)}$$

*and define*

$$D_{c,\lambda} := \{x \in \mathbb{R} \mid a_s \Lambda_x \in K_{r_{c,\lambda}(s)} \text{ for an unbounded set of } s\}.$$

*If the series (4-3) diverges, then the set*

$$D := \bigcap_{c\in\mathbb{R}} \bigcap_{\lambda>0} D_{c,\lambda}$$

*has full Lebesgue measure.*

**Remark 4.5.** We note that by our assumption $r_{c,\lambda}(\cdot)$ is not necessarily always positive, and the set $K_{r_{c,\lambda}(s)}$ is empty whenever $r_{c,\lambda}(s)$ is negative.

*Proof of Lemma 4.4.* For any function $f : [s_f, \infty) \to (0, \infty)$ with $s_f \geq 1$ we define

$$A_{\infty,f} := \{x \in \mathbb{R} \mid a_s \Lambda_x \in K_{f(s)} \text{ for an unbounded set of } s > s_f\}$$

and

$$N_f := \sum_{n\geq s_f} f(n) \log\left(\frac{1}{f(n)}\right).$$

First we note that the divergence of the series $N_r$ is equivalent to the divergence of the series $N_{r_c/2}$ for any $c \in \mathbb{R}$, where $r_c(s) := r(s + c) = r_{c,0}(s)$. Moreover, it is clear that $(r_c/2)(\cdot)$ satisfies the assumptions in Proposition 4.3. Thus, by Proposition 4.3, if the series $N_r$ diverges, then the set $A_{\infty,r_c/2}$ is of full Lebesgue measure for any $c \in \mathbb{R}$. On the other hand, for any $c \in \mathbb{R}$ and $\lambda > 0$ let $f_{c,\lambda}(s) = \lambda e^{-2(s+c)}$. It is easy to check that $f_{c,\lambda}|_{[s_{c,\lambda},\infty)}$ satisfies the assumptions in Proposition 4.3 with

$$s_{c,\lambda} := \max\left\{\frac{\log(2\lambda)}{2} - c, 1\right\},$$

and the series $N_{f_{c,\lambda}}$ converges for any $c \in \mathbb{R}$ and $\lambda > 0$. Thus by Proposition 4.3 the set $A_{\infty,f_{c,\lambda}}$ is of zero Lebesgue measure for any $c \in \mathbb{R}$ and $\lambda > 0$. Define

$$\bar{A} := \bigcap_{c\in\mathbb{R}} A_{\infty,r_c/2} \quad \text{and} \quad \underline{A} := \bigcup_{c\in\mathbb{R}} \bigcup_{\lambda>0} A_{\infty,f_{c,\lambda}}.$$

We note that since $r(\cdot)$ is nonincreasing, for any $c_1 < c_2$ we have $r_{c_1}/2 \geq r_{c_2}/2$ implying $A_{\infty,c_2/2} \subset A_{\infty,r_{c_1}/2}$. Hence the family of sets $\{A_{\infty,r_c/2}\}_{c\in\mathbb{R}}$ is nested and $\bar{A} = \lim_{c\to\infty} A_{\infty,r_c/2}$ is of full Lebesgue measure. Similarly, the family of sets $\{A_{\infty,f_{c,\lambda}}\}_{c\in\mathbb{R},\lambda>0}$ is also nested and the set

$$\underline{A} = \lim_{c\to-\infty}\lim_{\lambda\to\infty} A_{\infty,f_{c,\lambda}}$$

is of zero Lebesgue measure. Thus the set $\bar{A} \smallsetminus \underline{A}$ is of full Lebesgue measure and it suffices to show that $\bar{A} \smallsetminus \underline{A} \subset D$. That is, for any $x \in \bar{A} \smallsetminus \underline{A}$ we want to show that for any $c \in \mathbb{R}$ and any $\lambda > 0$ the events $a_s\Lambda_x \in K_{r_{c,\lambda}(s)}$ happen for an unbounded set of $s$. First we note that $x \in \bar{A}$ means that for any $c \in \mathbb{R}$ there exists an unbounded subset $S_c \subset \mathbb{R}$ such that $a_s\Lambda_x \in K_{r_c(s)/2}$ for any $s \in S_c$. Secondly, we note that $x \notin \underline{A}$ means that for any $c \in \mathbb{R}$ and $\lambda > 0$ there exists some constant $T_{c,\lambda} > 0$ such that for any $s \geq T_{c,\lambda}$ we have $a_s\Lambda_x \in \Delta^{-1}(f_{c,\lambda}(s), \infty)$. In particular, for any $s \in S_c \cap (T_{c,\lambda}, \infty)$ we have

$$f_{c,\lambda}(s) < \Delta(a_s\Lambda_x) \leq \frac{r_c(s)}{2}.$$

This implies

$$0 < \Delta(a_s\Lambda_x) \leq \frac{r_c(s)}{2} < \frac{r_c(s)}{2} + \frac{r_c(s)}{2} - f_{c,\lambda}(s) = r_{c,\lambda}(s)$$

for any $s \in S_c \cap (T_{c,\lambda}, \infty)$. Finally, we finish the proof by noting that since $S_c$ is unbounded, the set $S_c \cap (T_{c,\lambda}, \infty)$ is also unbounded. □

We can now give:

*Proof of Theorem 1.1.* The convergent case follows directly from Corollary 1.4 and the classical Borel–Cantelli lemma, and we thus only need to prove the divergent case. Let $r : [s_0, \infty) \to (0, \infty)$ be continuous, nonincreasing, satisfying (4-1) and such that the series (4-3) diverges; we want to show that $\mu(B_\infty) = 1$. First we note that we can assume $\lim_{s\to\infty} r(s) = 0$, since otherwise the result would follow from the ergodicity of the flow $\{a_s\}_{s>0}$ on $X$. Let $D := \bigcap_{c\in\mathbb{R}}\bigcap_{\lambda>0} D_{c,\lambda}$ be as in Lemma 4.4 and define $B \subset X$ such that

$$B = \left\{ \begin{pmatrix} a & 0 \\ b & a^{-1} \end{pmatrix} \Lambda_x \in X \;\middle|\; b \in \mathbb{R},\ a > 0,\ x \in D \right\}.$$

We note that by Lemma 4.4 the set $D$ has full Lebesgue measure. Thus the set $B \subset X$ is also of full measure (with respect to $\mu$) and it suffices to show that $B \subset B_\infty$. First, by direct computation for

$$\Lambda = \begin{pmatrix} a & 0 \\ b & a^{-1} \end{pmatrix} \Lambda_x \in B$$

we have

$$a_s\Lambda = \begin{pmatrix} 1 & 0 \\ e^{-2s}a^{-1}b & 1 \end{pmatrix} a_{s+\log a}\Lambda_x. \tag{4-6}$$

Next, for any $y \in \mathbb{R}$ let

$$u_y^- = \begin{pmatrix} 1 & 0 \\ y & 1 \end{pmatrix}.$$

Note that for any $\boldsymbol{v} \in \mathbb{R}^2$, we have $\|u_y^- \boldsymbol{v}\| \le (|y|+1)\|\boldsymbol{v}\|$. This implies that for any $\Lambda \in X$

$$|\Delta(u_y^- \Lambda) - \Delta(\Lambda)| \le \log(1+|y|).$$

Using the above inequality, the relation (4-6), and the inequality $\log(1+x) < 2x$ for all $x > 0$, we get

$$|\Delta(a_s \Lambda) - \Delta(a_{s+\log a} \Lambda_x)| \le 2a^{-1}|b|e^{-2s}.$$

Since $x \in D$ for any $c \in \mathbb{R}$ and any $\lambda > 0$ we have $a_s \Lambda_x \in K_{r_{c,\lambda}(s)}$ for an unbounded set of $s$. In particular, taking $c = -\log a$, $\lambda = 2a^{-1}|b|$ we get

$$0 \le \Delta(a_s \Lambda) \le \Delta(a_{s-c} \Lambda_x) + \lambda e^{-2s} \le r_{c,\lambda}(s-c) + \lambda e^{-2s} = r(s)$$

for an unbounded set of $s$, finishing the proof. □

## Acknowledgements

The authors would like to thank Anurag Rao, Nick Wadleigh and Cheng Zheng for many helpful conversations. Thanks are also due to the anonymous referee for a quick and careful report.

## References

[Athreya 2009] J. S. Athreya, "Logarithm laws and shrinking target properties", *Proc. Indian Acad. Sci. Math. Sci.* **119**:4 (2009), 541–557. MR Zbl

[Athreya and Konstantoulas 2016] J. S. Athreya and I. Konstantoulas, "Discrepancy of general symplectic lattices", preprint, 2016. arXiv

[Athreya and Margulis 2009] J. S. Athreya and G. A. Margulis, "Logarithm laws for unipotent flows, I", *J. Mod. Dyn.* **3**:3 (2009), 359–378. MR Zbl

[Cassels 1997] J. W. S. Cassels, *An introduction to the geometry of numbers*, Springer, 1997. MR Zbl

[Chernov and Kleinbock 2001] N. Chernov and D. Kleinbock, "Dynamical Borel–Cantelli lemmas for Gibbs measures", *Israel J. Math.* **122** (2001), 1–27. MR Zbl

[Dani 1985] S. G. Dani, "Divergent trajectories of flows on homogeneous spaces and Diophantine approximation", *J. Reine Angew. Math.* **359** (1985), 55–89. MR Zbl

[Davenport and Schmidt 1970a] H. Davenport and W. M. Schmidt, "Dirichlet's theorem on diophantine approximation", pp. 113–132 in *Symposia Mathematica, IV* (Rome, 1968/69), Academic Press, London, 1970. MR Zbl

[Davenport and Schmidt 1970b] H. Davenport and W. M. Schmidt, "Dirichlet's theorem on diophantine approximation, II", *Acta Arith.* **16** (1970), 413–424. MR Zbl

[Fairchild 2019] S. K. Fairchild, "A higher moment formula for the Siegel–Veech transform over quotients by Hecke triangle groups", preprint, 2019. arXiv

[Kelmer and Yu 2019] D. Kelmer and S. Yu, "The second moment of the Siegel transform in the space of symplectic lattices", *Int. Math. Res. Not.* (online publication February 2019).

[Kleinbock and Margulis 1999] D. Y. Kleinbock and G. A. Margulis, "Logarithm laws for flows on homogeneous spaces", *Invent. Math.* **138**:3 (1999), 451–494. Correction in **211**:2 (2018), 855–862. MR Zbl

[Kleinbock and Rao 2019] D. Kleinbock and A. Rao, "A zero-one law for uniform Diophantine approximation in Euclidean norm", preprint, 2019. arXiv

[Kleinbock and Wadleigh 2018] D. Kleinbock and N. Wadleigh, "A zero-one law for improvements to Dirichlet's theorem", *Proc. Amer. Math. Soc.* **146**:5 (2018), 1833–1844. MR Zbl

[Lang 1975] S. Lang, $\mathrm{SL}_2(\mathbb{R})$, Addison-Wesley, Reading, MA, 1975. MR Zbl

[Maucourant 2006] F. Maucourant, "Dynamical Borel–Cantelli lemma for hyperbolic spaces", *Israel J. Math.* **152** (2006), 143–155. MR Zbl

[Randol 1970] B. Randol, "A group-theoretic lattice-point problem", pp. 291–295 in *Problems in analysis* (*papers dedicated to Salomon Bochner, 1969*), 1970. MR Zbl

[Rogers 1955] C. A. Rogers, "Mean values over the space of lattices", *Acta Math.* **94** (1955), 249–287. MR Zbl

[Schmidt 1960] W. M. Schmidt, "A metrical theorem in geometry of numbers", *Trans. Amer. Math. Soc.* **95** (1960), 516–529. MR Zbl

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

[Siegel 1945] C. L. Siegel, "A mean value theorem in geometry of numbers", *Ann. of Math.* (2) **46** (1945), 340–347. MR Zbl

[Sullivan 1982] D. Sullivan, "Disjoint spheres, approximation by imaginary quadratic numbers, and the logarithm law for geodesics", *Acta Math.* **149**:3-4 (1982), 215–237. MR Zbl

[Veech 1998] W. A. Veech, "Siegel measures", *Ann. of Math.* (2) **148**:3 (1998), 895–944. MR Zbl

[Wood 1992] D. C. Wood, "The computation of polylogarithms", technical report 15-92*, University of Kent, Computing Laborabory, 1992, available at https://www.cs.kent.ac.uk/pubs/1992/110/content.pdf.

Received 2 Oct 2019. Revised 30 Dec 2019.

DMITRY KLEINBOCK:

kleinboc@brandeis.edu
Department of Mathematics, Brandeis University, Waltham, MA, United States

SHUCHENG YU:

yushucheng@campus.technion.ac.il
Department of Mathematics, Technion, Haifa, Israel

msp

Moscow Journal of Combinatorics and Number Theory
Vol. 9, No. 2, 2020
https://doi.org/10.2140/moscow.2020.9.123

# Algebraic cryptanalysis and new security enhancements

Vitaliĭ Roman'kov

We briefly discuss linear decomposition and nonlinear decomposition attacks using polynomial-time deterministic algorithms that recover the secret shared keys from public data in many schemes of algebraic cryptography. We show that in this case, contrary to common opinion, typical computational security assumptions are not very relevant to the security of the schemes; i.e., one can break the schemes without solving the algorithmic problems on which the assumptions are based. Also we present another and in some points similar approach, which was established by Tsaban et al.

Before demonstrating the applicability of these two methods to two well-known noncommutative protocols, we cryptanalyze two new cryptographic schemes that have not yet been analyzed.

Further, we introduce a novel method of construction of systems resistant against attacks via linear algebra. In particular, we propose improved versions of the well-known Diffie–Hellman-type (DH) and Anshel–Anshel–Goldfeld (AAG) algebraic cryptographic key-exchange protocols.

## 1. Introduction

In [Roman'kov 2013a], the author introduced a method of *linear decomposition* applicable in algebraic cryptanalysis. This method was further developed in [Myasnikov and Roman'kov 2015]; see also [Roman'kov 2013b; 2018a; 2018b]. In [Roman'kov 2016], this method was supplemented by a *nonlinear decomposition* method; see also [Roman'kov 2018b]. These methods can be applied for obtaining secret keys without computing private parameters or solving algorithmic problems on which the protocols are based. These applications are called *linear* and *nonlinear decomposition attacks* respectively. They are deterministic, provable and polynomial-time. These methods were widely applied in cryptanalysis of dozens of protocols of algebraic cryptography; see [Roman'kov 2018b]. The linear decomposition attack can be applied to protocols based on matrix groups over arbitrary (finite or infinite) fields. The nonlinear decomposition attack is applicable to protocols based on groups that are not necessary matrix, or do not use matrix representations. See details in [Roman'kov 2016; 2018b].

B. Tsaban [2015] introduced a method for obtaining provable polynomial-time solutions of problems in noncommutative algebraic cryptography called the *linear span-method*, or simply the *span-method*; see also [Ben-Zvi et al. 2018]. This method is probabilistic and is a fundamental base for algebraic span cryptanalysis, a general approach for provable polynomial-time solutions of computational problems in groups of matrices over finite fields, and thus in all groups with efficient matrix representations over finite fields. This approach is widely applicable; in particular, it is applicable to the AAG protocol. Algebraic

This work was supported by the Mathematical Center in Akademgorodok under agreement no. 075-15-2019-1613 with the Ministry of Science and Higher Education of the Russian Federation.

*MSC2010:* primary 20F10; secondary 20F70, 94A60.

*Keywords:* postquantum cryptography, algebraic cryptanalysis, algebraic cryptography, marginal sets.

span cryptanalysis improves upon earlier approaches, such as Cheon–Jun's method [2003] and Tsaban's linear centralizer method [2015].

We will not describe these methods in detail, but we will give a couple of examples of how these methods can be applied. Some of these applications, namely to the DH and to the AAG protocols, were previously presented in the literature. We present them here because we propose improved versions of them. There are exactly two new applications: one of them to cryptanalysis of the ElGamal-type version of the cryptosystem MOR introduced in [Bhunia et al. 2019], and the other to the cryptosystem proposed in [Baba et al. 2011].

A different probabilistic attack on the braid group cryptosystems is the length-based attack. The length-based attack on AAG protocol was initially proposed by J. Hughes and A. Tannenbaum [2002]. A. D. Myasnikov and A. Ushakov [2007] showed that accurately designed length-based attack can successfully break a random instance of the simultaneous conjugacy search problem for certain parameter values and argued that the public/private information chosen uniformly random leads to weak keys. This attack can be applied to other groups too. See [Garber et al. 2006; Hofheinz and Steinwandt 2002; Hughes 2002; Myasnikov et al. 2005; 2006].

The presence of effective methods of linear algebra in algebraic cryptanalysis requires the development of tools to counter these methods. Section 7 presents such tools. Their use makes some well-known schemes protected against attacks by the linear algebra methods. As examples of such protection, we provide improved versions of the DH and AAG algebraic cryptographic key-exchange protocols.

Throughout we use the following notation:

- $\mathbb{Z}$, the set of integer numbers.
- $\mathbb{N}$, the set of nonnegative integer numbers.
- $\mathbb{S}_n$, the symmetric group of degree $n$.
- $g^h = hgh^{-1}$, conjugate.
- $[g, h] = ghg^{-1}h^{-1}$, commutator.

For a group $G$, we have:

- $G'$, commutant (derived subgroup).
- $\mathrm{C}_G(A)$, centralizer of $A$ in $G$.
- $\mathrm{Aut}(G)$, automorphism group.

## 2. Mathematical background for the linear algebra methods

Let $\mathbb{F}$ be a field and $\mathrm{M}(n, \mathbb{F})$ be the set of $n \times n$ matrices with entries in $\mathbb{F}$. For a set $S \subseteq \mathrm{M}(n, \mathbb{F})$, let $\mathrm{Alg}(S)$ be the algebra generated by $S$, that is, the smallest algebra $A \subseteq \mathrm{M}(n, \mathbb{F})$ that contains $S$ as a subset. Every subalgebra of $\mathrm{M}(n, \mathbb{F})$ is also a vector space over the field $\mathbb{F}$. Let $\mathrm{GL}(n, \mathbb{F})$ be the group of invertible matrices in $\mathrm{M}(n, \mathbb{F})$. For a subgroup $G \leq \mathrm{GL}(n, \mathbb{F})$, we have $\mathrm{Alg}(G) = \mathrm{span}(G)$, where $\mathrm{span}(G)$ is the vector space spanned by $G$.

**Proposition 2.1** [Ben-Zvi et al. 2018, Proposition 1]. *Let $G = \mathrm{gp}(g_1, \ldots, g_k) \leq \mathrm{GL}(n, \mathbb{F})$ be a group, and $d \leq n^2$ be the dimension of the vector space $\mathrm{Alg}(G)$. A basis for the vector space $\mathrm{Alg}(G)$ can be computed using $O(kd^2n^2)$ field operations.*

**Lemma 2.2** (invertibility lemma [Tsaban 2015, Lemma 9]). *For a finite field $\mathbb{F}_q$ of order $q$, let $h_1, \dots, h_m \in \mathrm{M}(n, \mathbb{F}_q)$ such that some linear combination of these matrices is invertible. If $\alpha_1, \dots, \alpha_m$ are chosen uniformly and independently from $\mathbb{F}_q$, then the probability that the linear combination $\alpha_1 h_1 + \cdots + \alpha_m h_m$ is invertible is at least $1 - n/q$.*

Let $V$ be a finite-dimensional vector space over a field $\mathbb{F}$ with basis $\mathcal{B} = \{v_1, \dots, v_r\}$. Let $\mathrm{End}(V)$ be the semigroup of endomorphisms of $V$. We assume that elements $v \in V$ are given as vectors relative to $\mathcal{B}$, and endomorphisms $a \in \mathrm{End}(V)$ are given by their matrices relative to $\mathcal{B}$. For an endomorphism $a \in \mathrm{End}(V)$ and an element $v \in V$ we denote by $v^a$ the image of $v$ under $a$. Also, for any subsets $W \subseteq V$ and $A \subseteq \mathrm{End}(V)$ we put $W^A = \{w^a : w \in W, a \in A\}$. We assume that elements of the field $\mathbb{F}$ are given in some constructive form and the "size" of the form is defined. Furthermore, we assume that the basic field operations in $\mathbb{F}$ are efficient; in particular they can be performed in polynomial time in the size of the elements. In other words, $F$ is *constructive*. For an element $\alpha \in \mathbb{F}$ we write $|\alpha|$ for the size of $\alpha$ and put $|v| = \max\{|\alpha_i|\}$ for a vector $v = (\alpha_1, \dots, \alpha_r) \in V$, and $|a| = \max\{|\alpha_{ij}|\}$ for a matrix $a = (\alpha_{ij}) \in \mathrm{End}(V)$.

**Lemma 2.3** (principal lemma [Myasnikov and Roman'kov 2015, Lemma 3.1]). *There is an algorithm that for given finite subsets $W \subseteq V$ and $U \subseteq \mathrm{End}(V)$ finds a basis of the subspace $\mathrm{span}(W^{\mathrm{sm}(U)})$ in the form $\{w_1^{a_1}, \dots, w_t^{a_t}\}$, where $w_i \in W$ and $a_i \in \mathrm{sm}(U)$. Here $\mathrm{sm}(U)$ denotes the submonoid generated by $U$. Furthermore, the number of field operations used by the algorithm is polynomial in $r = \dim(V)$ and the cardinalities of $W$ and $U$. The total estimate is $O(r^3|U|^2 + r|W|^2)$.*

## 3. Cryptanalysis of two schemes of Baba et al. by the linear algebra methods

In [Baba et al. 2011], S. Baba, S. Kotyada and R. Teja demonstrated how to define a supposedly one-way function FACTOR in a noncommutative group. As an example of a platform for implementing FACTOR, they proposed one of the groups, such as $\mathrm{GL}(n, \mathbb{F}_q)$, $\mathrm{UT}(n, \mathbb{F}_q)$ or braid groups $B_n$, $n \in \mathbb{N}$. Here $\mathbb{F}_q$ denotes a finite field of order $q$.

They believed that the function FACTOR was one-way, which means that the inverse to FACTOR is easy to compute, while the function itself is hard to compute. Shortly afterwards Stanek [2011] published an extension of the baby-step giant-step algorithm disproving this conjecture. Note that the baby-step giant-step methods are limited in practice because of memory requirements. In [Romsy 2011] a modification of Pollard's kangaroo algorithm was presented that solves the FACTOR problem requiring only negligible memory. Anyway these methods have very complicated implementations. We will show that the linear algebra approach is much simpler and more efficient. At the same time, this will be an example of using the methods presented.

Then, using the FACTOR function as a primitive, the authors of [Baba et al. 2011] defined a public key cryptosystem which is comparable to the classical ElGamal system based on the discrete logarithm problem. Recall, that the ElGamal system can be described as follows: Let $G$ be a public finite cyclic group with generator $g$, and let $x \in \mathbb{Z}$ be Alice's private key. The element $g^x$ is public. To send a message $m \in G$, Bob picks a random integer $y$ and sends the ciphertext $c = (g^y, g^{xy}m)$ to Alice. To decrypt, Alice calculates $(g^y)^x = g^{xy}$ and inverts it to retrieve $m$. There are a couple of cryptosystems of ElGamal-type. See, for example, [Kahrobaei and Khan 2006; Fine et al. 2016]. The versions proposed in [Mahalanobis 2008; 2012] were analyzed in [Roman'kov and Obzor 2018]. See also cryptanalysis in [Roman'kov 2018b].

In [Baba et al. 2011], the authors also proposed a key exchange, analogous to the DH key exchange protocol in a noncommutative setting using FACTOR. Recall, that the classical DH protocol can be described as follows: Let $G$ be a public finite cyclic group with generator $g$, and let $x \in \mathbb{Z}$ be Alice's private key and $y \in \mathbb{Z}$ be Bob's private key. Alice publishes $g^x$ and Bob publishes $g^y$. Then each of them computes the exchanged key $g^{xy} = (g^x)^y = (g^y)^x$.

In this paper, we apply and compare two methods of algebraic cryptanalysis via linear algebra, namely, the linear decomposition method invented and developed by the author in [Roman'kov 2013a; 2013b; 2018b] and in [Myasnikov and Roman'kov 2015], and the span-method invented by B. Tsaban and developed with A. Ben-Zvi, and A. Kalka [Tsaban 2015; Ben-Zvi et al. 2018] to show the vulnerability of the cryptosystem and protocol proposed in [Baba et al. 2011].

**3A.** ***The ElGamal-type cryptosystem based on FACTOR [Baba et al. 2011].*** Let $G$ be any group and let $g, h \in G$ be two noncommuting elements chosen by Alice. Let $\mathrm{gp}(g)$ and $\mathrm{gp}(h)$ be the cyclic subgroups generated by these elements, respectively. In order to define the FACTOR function one assume that $\mathrm{gp}(g) \cap \mathrm{gp}(h) = \{1\}$. Let $\varphi : \mathrm{gp}(g) \times \mathrm{gp}(h) \to G$ be a function defined by $\varphi(g^x, h^y) = g^x \cdot h^y$, where $x, y \in \mathbb{Z}$. Obviously, $\varphi$ is injective. Then $\mathrm{FACTOR}(g^x h^y) = \varphi^{-1}(g^x h^y)$.

We suppose that Alice is the recipient of the messages and Bob is communicating with Alice. Let $m \in G$ be a message.

**Algorithm.**

- Alice picks arbitrary random integers $x, y \in \mathbb{Z}$ and sets a public key $(G, g, h, g^x h^y)$. Alice has a private key $(g^x, h^y)$ for decryption.
- To send $m$, Bob picks arbitrary random private integers $x', y'$ and sends the ciphertext

$$c = (g^{x+x'} h^{y+y'}, g^{x'} h^{y'} m)$$

  to Alice.
- To decrypt the ciphertext, Alice uses her private key and calculates

$$(g^x)^{-1}(g^{x+x'} h^{y+y'})(h^y)^{-1} = g^{x'} h^{y'}.$$

  Then she inverts it to retrieve $m$.

The authors of this scheme hoped that the security of the cryptosystem described above reduces to solving FACTOR problem in the underlying group. Below we will show that the system is vulnerable to linear algebra attacks.

**3B.** ***Cryptanalysis of the ElGamal-type cryptosystem based on FACTOR.*** We will show that any intruder can efficiently retrieve $m$.

First we will use the span-method.

**Theorem 3.1.** *Suppose that $G$ is a finite group presented as a matrix group over a finite field $\mathbb{F}_q$ of order $q$; i.e., $G \leq \mathrm{M}(n, \mathbb{F}_q)$. Let $g, h \in G$ be two noncommuting elements such that $\mathrm{gp}(g) \cap \mathrm{gp}(h) = \{1\}$. Given $g^x h^y$, $g^{x+x'} h^{y+y'} \in G$, where $x, x', y, y' \in \mathbb{N}$, one can find in polynomial time (in the size of the public data) the element $g^{x'} h^{y'}$.*

*Proof.* Let $V = \mathrm{span}(\mathrm{gp}(g))$ be the linear subspace of $\mathrm{M}(n, \mathbb{F}_q)$ generated by all matrices of the form $g^i$, $i \in \mathbb{Z}$. Then $\dim(V) \le n-1$. Since $g$ is the root of its characteristic polynomial of degree $n$, matrices $1, g, g^2, \ldots, g^n$ are linearly dependent. Obviously, if $g^{k+1}$ lies in $\mathrm{span}(\{1, g, g^2, \ldots, g^k\})$, then $g^{k+t}, g^{1-t} \in \mathrm{span}(\{1, g^2, \ldots, g^k\})$ for every $t = 2, 3, \ldots$.

By the Gaussian elimination method, we can efficiently construct a basis for $V$. For example, we can take as a basis the maximum independent set of elements of the form $1, g, g^2, \ldots, g^k$, checking for each subsequent $l = 0, 1, \ldots$ whether or not $g^{l+1}$ lies in $\mathrm{span}(\{1, g, g^2, \ldots, g^l\})$.

Consider the equation

$$f(g^x h^y)h = hf(g^x h^y) \sim fg^x h = hfg^x, \tag{1}$$

which is linear with respect to $n^2$ unknown entries of matrix $f$. We will seek $f$ in the form

$$f = \sum_{i=0}^{k} \alpha_i g^i;$$

i.e., we seek a solution $f$ in $V$. We know that there is a nondegenerate solution $f = g^{-x}$. By Proposition 2.1 we can efficiently construct a basis $e_1, \ldots, e_p$ of the subspace of all solutions of (1) in $V$. Then we apply the invertibility lemma, Lemma 2.2 to find an invertible solution $f$.

Let the element $f$ be found. Then

$$f(g^{x+x'}h^{y+y'}) = g^{x'}(f(g^x h^{y'}))h^y = (g^{x'}h^{y'})f(g^x h^y)$$

and

$$(g^{x'}h^{y'})f(g^x h^y)(g^x h^y)^{-1} f^{-1} = g^{x'}h^{y'}. \qquad \square$$

Now we apply the result just obtained to the protocol under consideration.

**Corollary 3.2.** *We have*

$$(g^{x'}h^{y'})^{-1}(g^{x'}h^{y'}m) = m,$$

*and the message $m$ is thus computed.*

Now we will show how, using the linear decomposition method, we can calculate the message $m$ for an arbitrary constructive field by a deterministic algorithm.

**Theorem 3.3.** *Let $G \le \mathrm{M}(n, \mathbb{F})$ be a matrix group over an arbitrary (constructive) field $\mathbb{F}$. Let $g, h \in G$ be two noncommuting elements such that $\mathrm{gp}(g) \cap \mathrm{gp}(h) = \{1\}$ and $m \in G$. Given the elements $g^x h^y$, $g^{x+x'}h^{y+y'}$, $g^{x+x'}h^{y+y'}m \in G$, where $x, x', y, y' \in \mathbb{Z}$, one can find in polynomial time (in the size of the public data) the element $m$.*

*Proof.* Let $V = \mathrm{span}(\mathrm{gp}(g)g^x h^y \mathrm{gp}(h))$ be the linear subspace of $\mathrm{M}(n, \mathbb{F})$ generated by all matrices of the form $g^i(g^x h^y)h^j$, $i, j \in \mathbb{Z}$. Then $\dim(V) \le (n-1)^2$. In the notation of Lemma 2.3, $V = W^{\mathrm{sm}}(U)$, $W = \{g^x h^y\}$, $U = \mathrm{sm}(l(g^{\pm 1}, r(h^{\pm 1}))$, where for any $f \in G$, $l(f)$ means the endomorphism of $\mathrm{M}(n, \mathbb{F})$ corresponding to left-sided multiplication by $f$. Similarly $r(f)$ means the endomorphism of $\mathrm{M}(n, \mathbb{F})$ corresponding to right-sided multiplication by $f$.

By Lemma 2.3, we can efficiently obtain a basis $e_i = g^{u_i}(g^x h^y)h^{v_i}$, $u_i, v_i \in \mathbb{Z}$, $i = 1, \ldots, r$ of $V$.

Since, $g^{x+x'}h^{y+y'} \in V$, by Lemma 2.3 we can efficiently obtain an expression of the form

$$g^{x+x'}h^{y+y'} = \sum_{i=1}^{r} \alpha_i e_i, \quad \alpha_i \in \mathbb{F},\ i = 1, \dots, r. \tag{2}$$

The right side of (2) is equal to

$$g^x \left( \sum_{i=1}^{r} \alpha_i g^{u_i} h^{v_i} \right) h^y, \tag{3}$$

and it follows by (2) that

$$g^{x'}h^{y'} = \sum_{i=1}^{r} \alpha_i g^{u_i} h^{v_i}. \tag{4}$$

The message $m$ is retrieved as above. □

**Remark 3.4.** Recall that the authors of [Baba et al. 2011] suggest as a platform for their cryptosystem one of the groups $\mathrm{GL}(n, \mathbb{F}_q)$, $\mathrm{UT}(n, \mathbb{F}_q)$, or braid groups $B_n$, $n \in \mathbb{N}$. In our cryptanalysis, we consider only matrix groups. Any group $B_n$ admits a faithful matrix representation [Bigelow 2001; Krammer 2002]. The braid group $B_n$ is linear via the so-called Lawrence–Krammer (LK) representation $B_n \to \mathrm{GL}(m, \mathbb{Z}[t^{\pm 1}, 1/2])$, where $m = n(n-1)/2$, which is injective. The LK representation can be computed by a polynomial-time algorithm. This representation is also invertible by (similar) polynomial-time algorithm; see [Krammer 2002; Cheon and Jun 2003].

**3C.** ***The DH key exchange protocol based on FACTOR [Baba et al. 2011] as a particular case of the protocol in [Sidelnikov et al. 1993].*** Suppose Alice and Bob want to exchange keys. Suppose $G, g, h$ are as in Section 3A.

**Algorithm 1.**
- Alice chooses a random pair of integers $(x_1, y_1)$. Then Alice sends the element $g^{x_1}h^{y_1}$ to Bob.
- Bob picks up two random integers $(x_2, y_2)$. Then Bob sends the element $g^{x_2}h^{y_2}$ to Alice.
- Alice computes $K_A = g^{x_1}(g^{x_2}h^{y_2})h^{y_1} = g^{x_1+x_2}h^{y_1+y_2}$.
- Bob computes $K_B = g^{x_2}(g^{x_1}h^{y_1})h^{y_2} = g^{x_1+x_2}h^{y_1+y_2}$.
- Now Alice and Bob have their exchanged secret key $K_1 = K_A = K_B$.

This algorithm is a particular case of the following algorithm of [Sidelnikov et al. 1993].
Let $G$ be a group, $A$ and $B$ two of its commutative subgroups, and $g \in G$. This data is public.

**Algorithm 2.**
- Alice chooses a random pair of elements $(a, b) \in A \times B$. Then Alice sends the element $agb$ to Bob.
- Bob picks up two random elements $(a', b') \in A \times B$. Then Bob sends the element $a'gb'$ to Alice.
- Alice computes $K_A = aa'gb'b$.
- Bob computes $K_B = a'agbb'$.
- Now Alice and Bob have their exchanged secret key $K_2 = K_A = K_B$.

**3D.** ***Cryptanalysis the DH key exchange protocols presented above.*** Now we will apply the linear decomposition method to reveal $K$.

**Theorem 3.5.** *Let $G \le \mathrm{M}(n, \mathbb{F})$ be a matrix group over an arbitrary constructive field $\mathbb{F}$. Let $g \in G$ and let $A = \mathrm{gp}(a_1, \dots, a_m)$, $B = \mathrm{gp}(b_1, \dots, b_s)$ be two finitely generated subgroups of $G$. Given $agb, a'gb'$, where $a, a' \in A,\ b, b' \in B$, one can find in polynomial time (in the size of the public data) the element $aa'gbb'$.*

*Proof.* Let $V = \mathrm{span}(AgB)$ be the linear subspace of $\mathrm{M}(n, \mathbb{F})$ generated by all matrices of the form $ugv$, $u \in A,\ v \in B$. Then $\dim(V) \le (n-1)^2$.

In the notation of Lemma 2.3, $V = W^{\mathrm{sm}}(U)$, where $W = \{g\},\ U = \mathrm{sm}(l(a_i^{\pm 1}), r(b_j \pm 1)),\ i = 1, \dots, m,\ j = 1, \dots, s$. Let $e_i = u_i g v_i\ \ i = 1, \dots, r$, be a basis of $V$ that can be efficiently obtained by Lemma 2.3

Since, $agb \in V$, we can efficiently obtain an expression of the form

$$agb = \sum_{i=1}^{r} \alpha_i e_i, \quad \alpha_i \in \mathbb{F},\ i = 1, \dots, r. \tag{5}$$

Then

$$\sum_{i=1}^{r} \alpha_i u_i (a'gb') v_i = a' \left( \sum_{i=1}^{r} \alpha_i e_i \right) b' = aa'gbb', \tag{6}$$

completing the proof. □

**Corollary 3.6.** *Each of the keys $K_1$ and $K_2$ of Algorithms 1 and 2 can be efficiently calculated in polynomial time (from the size of the public data of the algorithms).*

The described cryptanalysis has many analogues, presented in [Roman'kov 2013a; 2013b; 2016; 2018a; 2018b; 2019a; Ben-Zvi et al. 2018; Tsaban 2015]. In [Roman'kov 2018a], a general scheme based on multiplications is presented. It corresponds to a number of cryptographic systems known in the literature, which are also vulnerable to attacks by the linear decomposition method. Note that Tsaban's span-method allows him to show the vulnerability of the well-known schemes of [Anshel et al. 1999], and the triple decomposition key exchange protocol of [Peker 2014].

## 4. Cryptanalysis of a new version of the MOR scheme

S. Bhunia, A. Mahalanobis, P. Shinde and A. Singh [Bhunia et al. 2019] studied the ElGamal-type version of the MOR cryptosystem with symplectic and orthogonal groups over finite fields $\mathbb{F}_q$ of odd characteristics. The MOR cryptosystem over $\mathrm{SL}(d, \mathbb{F}_q)$ was previously investigated by the second of these authors. In that case, the hardness of the MOR cryptosystem was found to be equivalent to the discrete logarithm problem in $F_{q^d}$. It is shown in [Bhunia et al. 2019] that the MOR cryptosystem over $\mathrm{Sp}(d, q)$ has the security of the discrete logarithm problem in $\mathbb{F}_{q^d}$. The MOR cryptosystem was also studied in [Paeng et al. 2001; Mahalanobis 2015] and was cryptanalyzed in [Monico 2016].

We are to show that the version of MOR in [Bhunia et al. 2019] is not entirely accurate. It should be supplemented with an additional assumption. The equivalence theorem there should be clarified too.

We also show that the proposed ElGamal-type version of MOR over any finitely generated matrix group $G \le \mathrm{GL}(d, \mathbb{F}_q)$ is vulnerable with respect to the linear decomposition attack in any case when the automorphism $\varphi$ can be naturally extended to a linear transformation of the linear space $\mathrm{span}(G)$

generated by $G$ in $\mathrm{M}(d, \mathbb{F})$, for example, if $\varphi$ is an inner automorphism. In fact, there exists an efficient algorithm to compute the original message by its ciphertext. It can be done for every constructive field, i.e., a field for which all operations are efficient, and the Gaussian elimination process is efficient too.

**4A.** ***The ElGamal version of the MOR cryptosystem [Bhunia et al. 2019].*** Let $G = \mathrm{gp}(g_1, g_2, \ldots, g_n)$ be a (finite) public group and $\varphi$ a nontrivial public automorphism of $G$.

Alice's keys are as follows:

- Private key: $t \in \mathbb{N}$.
- Public key: $\{\varphi(g_i) : i = 1, \ldots, n\}$ and $\{\varphi^t(g_i) : i = 1, \ldots, n\}$.

We suppose that Alice is the recipient of the messages and Bob is communicating with Alice. Let $m \in G$ be a message.

**Algorithm.**
- To send the message (plaintext) $m$ Bob picks up a random integer $r$, then he computes $\{\varphi^r(g_i) : i = 1, \ldots, n\}$ and $\varphi^{tr}(m)$. The ciphertext is $(\{\varphi^r(g_i) : i = 1, \ldots, n\}, \varphi^{tr}(m))$.
- Since Alice knows $t$, she computes $\varphi^{tr}(g_i)$ from $\varphi^r(g_i)$ and then $\varphi^{-tr}(g_i)$, $i = 1, \ldots, n$. Finally, the message $m$ can be computed by $\varphi^{-tr}(\varphi^{tr}(m)) = m$.

**Remark 4.1.** There is one obstacle to the implementation of the decryption process. To recover $m$, Alice should compute $\{\varphi^{-tr}(g_i) : i = 1, , n\}$ by $\{\varphi^{tr}(g_i) : i = 1, \ldots, n\}$ or compute it by $\varphi^r$. It can be done if she knows $\varphi^{-1}$, i.e., $\{\varphi^{-1}(g_i) : i = 1, \ldots, n\}$.

In the general case, the calculation of the inverse automorphism is not an obviously efficient process. We have to assume that Alice can do it, for example, because she knows $s \in \mathbb{N}$ such that $\varphi^s = \mathrm{id}$. It happens, in particular, if she knows the order $s_1$ of $\varphi$ or the order $s_2$ of $\mathrm{Aut}(G)$. Then $\varphi^{-1} = \varphi^{s-1}$ ($s = s_1$ or $s = s_2$). Also Alice can know $\varphi^{-1}$.

Alice can simultaneously build $\varphi$ and $\varphi^{-1}$ during the setting of parameters of the protocol.

This obstacle manifests itself more significantly in the proof of the following theorem.

**Theorem** [Bhunia et al. 2019, Theorem 2.1]. *The difficulty in breaking the above MOR cryptosystem is equivalent to the DH problem in the group* $\mathrm{gp}(\varphi)$.

*Proof.* It is easy to see that if one can break the DH problem, then one can compute $\varphi^{tr}$ from $\varphi^t$ in the public key and $\varphi^r$ in the ciphertext. This breaks the system.

On the other hand, observe that the plaintext is $m = \varphi^{-tr}(\varphi^{tr}(m))$. Assume that there is an oracle that can break the MOR cryptosystem, i.e., given $\varphi, \varphi^t$ and a ciphertext $(\varphi^r, f)$ will deliver $\varphi^{-tr}(f)$. Now we query the oracle $n$ times with the public key and the ciphertexts $(\varphi^r(g_i), g_i)$ for $i = 1, \ldots, n$. From the output, one can easily find $\varphi^{-tr}(g_i)$ for $i = 1, 2, \ldots, n$. So we just witnessed that for $\varphi^r(g_i)$ and $\varphi^t(g_i)$ for $i = 1, \ldots, n$, one can compute $\varphi^{-tr}(f)$ for every $f = f(g_1, \ldots, g_n)$ using the oracle. This solves the DH problem. □

**Remark 4.2.** In the first part of the proof one computes $\varphi^{tr}$, but one needs $\varphi^{-tr}$ to compute $m$ in the protocol. However, it is not always easy to find the inverse. There are some cryptographic schemes based on the complexity of the problem of finding the inverse to a given automorphism.

**4B.** ***Cryptanalysis of the ElGamal version of the MOR cryptosystem.*** We propose the following cryptanalysis that works in the case of an arbitrary (constructive) field $\mathbb{F}$.

Suppose that the ElGamal-type system MOR is considered over a finitely generated matrix group $G \leq \mathrm{GL}(d, \mathbb{F})$. Then $G \subseteq \mathrm{M}(d, \mathbb{F})$. Let $G = \mathrm{gp}(g_1, \ldots, g_n)$. We suppose that $\varphi$ can be naturally extended to a linear transformation of $V = \mathrm{span}(G)$ that is a linear subspace generated by $G$ in $\mathrm{M}(d, \mathbb{F})$. It happens for example, if $\varphi$ is an inner automorphism of $G$. Note, that the case of inner automorphism $\varphi$ is considered in [Bhunia et al. 2019] as the most significant.

To reveal $m$ using only open protocol data, we perform the following actions.

Step 1: Let $V_i,\ i \in \{1, \ldots, n\}$, be the subspace of $V$ generated by all elements of the form $\varphi^k(g_i)$ for $k \in \mathbb{Z}$. There is a basis of $V_i$ of the form $e_1(i) = \varphi^0(g_i) = g_i,\ e_2(i) = \varphi(g_i), \ldots, e_{l_i}(i) = \varphi^{l_i-1}(g_i)$. It can be efficiently constructed as follows.

Initially, we include $e_1(i) = g_i$ in the basis. Then we check whether $\varphi(g_i)$ belongs to the linear subspace generated by $e_1(i)$. If not, then we add $e_2(i) = \varphi(g_i)$ to the basis under construction. Suppose $e_1(i), \ldots, e_j(i)$ is a constructed part of the basis. Then we check whether $\varphi^j(g_i) = \varphi(e_j(i))$ belongs to the linear subspace generated by $e_1(i), \ldots, e_j(i)$. If not, then we add $e_{j+1}(i) = \varphi^j(g_i)$ to the basis under construction, and continue. If so, we stop the process and claim that the basis is constructed and $l_i = j$. Indeed, a linear presentation of $\varphi^j(g_i)$ via $e_1(i), \ldots, e_j(i)$ after applying $\varphi$ gives a linear presentation of $\varphi^{j+1}(g_i)$ via $e_2(i), \ldots, e_j(i), \varphi^j(g_i)$, and so via $e_1(i), \ldots, e_j(i)$. This argument works for every $j + v$, $v \geq 1$. Similarly we can obtain the linear decomposition of each $\varphi^{-v}(g_i),\ v \geq 1$.

Step 2: For each $i = 1, \ldots, n$, we have constructed a basis $e_1(i), \ldots, e_{l_i}(i)$ of $V_i$, where $e_{j+1}(i) = \varphi^j(g_i)$, $j = 0, \ldots, l_i - 1$. Each subspace $V_i$ is $\varphi$-invariant. In the general case, $l_i \leq d^2$.

In [Bhunia et al. 2019], the authors single out as the main the case of inner automorphism $\varphi$. They write:

> The purpose of this section is to show that for a secure MOR cryptosystem over the classical Chevalley and twisted orthogonal groups, we have to look at automorphisms that act by conjugation like the inner automorphisms. There are other automorphisms that also act by conjugation, like the diagonal automorphism and the graph automorphism for odd-order orthogonal groups. Then we argue what is the hardness of our security assumptions.

Then they note that by the Dieudonné theorem, $\varphi = \sigma\iota\eta\gamma\theta$, where $\sigma$ is a central automorphism, $\iota$ is an inner automorphism, $\eta$ is a diagonal automorphism, $\gamma$ is a graph automorphism, and $\theta$ is a field automorphism.

Then they continue:

> The group of central automorphisms is too small and the field automorphisms reduce to a discrete logarithm in the field $F_q$. So there is no benefit of using these in a MOR cryptosystem. Also there are not many graph automorphisms in classical Chevalley and twisted orthogonal groups other than special linear groups and odd-order orthogonal groups. In the odd-order orthogonal groups, these automorphisms act by conjugation.
>
> Recall that our automorphisms are presented as actions on generators. It is clear [Mahalanobis 2012, Section 7] that if we can recover the conjugating matrix from the action on the generators, the security is a discrete logarithm problem in $\mathbb{F}_{q^d}$, or else the security is a discrete logarithm problem in $F_{q^{d^2}}$.

In our cryptanalysis, we assume that $\varphi$ can be naturally extended to an automorphism of the linear space $V$. This happens if $\varphi$ is an inner or field automorphism or is induced by an inner automorphism of $\mathrm{GL}(d, \mathbb{F})$.

We return to the above-introduced subspaces $V_i$, $i = 1, \dots, n$. For a fixed $V_i$, denote by $\varphi_i$ the linear map of $V_i$ induced by $\varphi$. The matrix $A(\varphi_i)$ in the basis $E_i = \{e_1(i), \dots, e_{l_i}(i)\}$ has the form

$$A(\varphi_i) = \begin{pmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ & & \ddots & \ddots & \ddots & \\ 0 & \cdots & \cdots & 0 & 1 & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & 1 \\ \alpha_1 & \alpha_2 & \cdots & \cdots & \cdots & \alpha_{l_i} \end{pmatrix},$$

where $\varphi(e_{l_i}(i)) = \sum_{k=1}^{l_i} \alpha_k e_k(i)$, $\alpha_k \in \mathbb{F}$.

In this way, we can efficiently compute for each $i$ the value $\varphi^{-1}(g_i)$ corresponding to the first row of $A(\varphi_i)^{-1}$. So we can compute $\varphi^{-1}$.

Now we know the matrices $A(\varphi_i)^{\pm 1}$, $A(\varphi_i)^{\pm r}$, $A(\varphi_i)^{\pm t}$, $i = 1, \dots, n$, and we need to calculate $r$ or $t$. Then we can calculate $\varphi_i^{-rt}$ and restore $m$. We can provide sufficient calculations using only one or more of the matrices above.

In [Menezes and Vanstone 1992], it was shown how the discrete logarithm problem in some special class of matrices can be reduced to the discrete logarithm problem in some extensions of the underlying field. In [Menezes and Wu 1997], these results were extended to show how the discrete logarithm problem in every group $\mathrm{GL}(d, \mathbb{F})$ can be reduced in probabilistic polynomial time to the similar problem in small extensions of $\mathbb{F}$. The case of a finitely generated nilpotent group is considered in [Roman'kov 2019b].

We see that matrix groups over finite fields offer no significant advantage for the implementation of cryptographic protocols whose security is based on the difficulty of computing discrete logarithms.

The described cryptanalysis has many analogues, presented in [Roman'kov 2013a; Myasnikov and Roman'kov 2015]. In [Roman'kov 2018a], a general scheme based on multiplications is presented. It corresponds to a number of cryptographic systems known in the literature, which are also vulnerable to attacks by the linear decomposition method. The nonlinear decomposition method was invented in [Roman'kov 2016]. The nonlinear method can be applied when the group chosen as the platform for a cryptographic scheme is not linear or the least degree of their representability by matrices is too big for efficient computations. See details in [Roman'kov 2018b].

A protection against linear algebra attacks was recently invented in [Roman'kov 2019a]. It is described in the case of the cryptographic scheme of [Anshel et al. 1999] but can be applied to the DH and some other schemes too. See [Roman'kov 2019c; 2019d]. Further, we'll present this protection in more detail. This version is improved with respect to [Anshel et al. 1999].

## 5. Cryptanalysis of the Ko et al. and Anshel–Anshel–Goldfeld classical protocols of algebraic cryptography

**5A.** ***Noncommutative analogues of the DH protocol.*** In algebraic cryptography, the following noncommutative analogues of the DH protocol are considered:

- An analogue with conjugations [Ko et al. 2000]: for a group $G$ and an element $g \in G$, determine by two elements $g^a = aga^{-1}$ and $g^b = bgb^{-1}$, where $a, b \in G$, $ab = ba$, the element $g^{ab} = abga^{-1}b^{-1} = bagb^{-1}a^{-1}$.
- An analogue with twoside multiplication: for a group $G$ and an element $g \in G$, determine by two elements of the form $aga'$ and $bgb'$, where $a, b \in G$, $ab = ba$, $a'b' = b'a'$, the element $abga'b' = bagb'a'$.
- An analogue with automorphisms: for a group $G$ and an element $g \in G$, determine by two elements of the form $\alpha(g)$ and $\beta(g)$, where $\alpha, \beta \in \mathrm{Aut}(G)$, $\alpha\beta = \beta\alpha$, the element $\alpha(\beta(g)) = \beta(\alpha(g))$.

The linear decomposition method under certain natural conditions into the group $G$ (first of all, this is the existence of an effective embedding in a finite-dimensional linear space) effectively solves each of these problems.

The case of two-sided multiplication in its slightly weak form was analyzed in Section 3C. Now we consider the case with conjugations. We will demonstrate two attacks, the first based on the linear decomposition, and the second based on the nonlinear decomposition.

**5B.** ***The Ko et al. protocol [2000].*** Let $G \le \mathrm{M}(n, \mathbb{F})$ be a public matrix group over an arbitrary (constructive) field $\mathbb{F}$, and let $g$ be a public element of $G$. Suppose that $A = \mathrm{gp}(a_1, \dots, a_k)$ and $B = \mathrm{gp}(b_1, \dots, b_l)$ are two pointwise commuting public subgroups of $G$.

Alice's keys are as follows:

- Private key: $a \in A$.
- Public key: $g^a = aga^{-1}$.

Bob's keys are as follows:

- Private key: $b \in B$.
- Public key: $g^b = bgb^{-1}$.

**Algorithm.**
- Alice sends $g^a$ to Bob.
- Bob sends $g^b$ to Alice.
- Since Alice knows $a$, she computes $(g^b)^a = g^{ab}$ from $g^b$.
- Since Bob knows $b$ he computes $(g^a)^b = g^{ba}$.
- Now both, Alice and Bob, know a secret key $K = g^{ab}$, because $ab = ba$.

**5C.** ***Cryptanalysis of the Ko et al. protocol.*** We will apply the linear and nonlinear decomposition attacks.

*Linear decomposition attack.* Let $V = \mathrm{span}(g^A)$ be the linear subspace of $\mathrm{M}(n, \mathbb{F})$ generated by all matrices of the form $g^c$, $c \in A$. Then $\dim(V) \le n^2$.

Let $e_1, e_2, \dots, e_r$ be a basis of $V$ that can be efficiently obtained; see [Roman'kov 2013a; 2018b; Myasnikov and Roman'kov 2015]. Let $e_i = g^{c_i}$, $c_i \in A$, $i = 1, \dots, r$.

Since, $g^a \in V$, we can efficiently obtain a presentation of the form

$$g^a = \sum_{i=1}^{r} \alpha_i e_i, \quad \alpha_i \in \mathbb{F}, \ i = 1, \dots, r. \tag{7}$$

Then

$$\sum_{i=1}^{r} \alpha_i g^{c_i} g^b g^{-c_i} = \left( \sum_{i=1}^{r} \alpha_i e_i \right)^b = K. \tag{8}$$

The exchanged key is recovered without computing the private parameters $a$ and $b$. We did not solve the underlined search conjugacy problem (to find $a$ by $g^a$ or to find $b$ by $g^b$).

*Nonlinear decomposition attack.* All assumptions and algorithms are the same as above except the assumption that $G$ is a linear group. In addition we suppose that every subgroup of $G$ is finitely generated and the membership problem for $G$ is efficiently decidable. For example, $G$ is a finitely generated nilpotent or more generally polycyclic group. See [Roman'kov 2016; 2018b] for details.

Let $g^A$ be subgroup of $G$ generated by all elements of the form $g^c, c \in A$. Let $g_i = g^{c_i},\ c_i \in A$, $i = 1, \ldots, r$, be a finite generating set of $g^A$. We suppose that this generating set can be efficiently constructed; see [Roman'kov 2016; 2018b] again.

Since, $g^a \in g^A$, we can efficiently obtain a presentation of the form

$$g^a = \prod_{i=1}^{s} g_{i_j}^{\epsilon_i}, \quad i_j \in \{1, \ldots, r\},\ \epsilon_i \in \{\pm 1\},\ i = 1, \ldots, s. \tag{9}$$

Then

$$\prod_{i=1}^{s} c_{i_j} (g^b)^{\epsilon_i} c_{i_j}^{-1} = \left( \prod_{i=1}^{s} g_{i_j}^{\epsilon_i} \right)^b = K. \tag{10}$$

The exchanged key is recovered without computing the private parameters $a$ and $b$. We did not solve the underlined search conjugacy problem (to find $a$ by $g^a$ or to find $b$ by $g^b$).

**5D.** ***The Anshel–Anshel–Goldfeld protocol [Anshel et al. 1999].*** M. Anshel, I. Anshel and D. Goldfeld [Anshel et al. 1999], see also [Myasnikov et al. 2008; 2011; Roman'kov 2012], proposed a group-based key exchange protocol that we call the AAG protocol. It works as follows.

Suppose two correspondents Alice and Bob want to exchange a key. They agree about a group $G$ given by a finite set of generators that is used as the platform. It is supposed that $G$ is equipped with an efficient normal form of its elements and the main group operations can be computed efficiently. All the information about $G$, the normal form and efficient algorithms to compute products of elements, its inversions and normal forms, is public. In particular, the word problem is efficiently solvable for $G$.

To exchange a key the correspondents act as follows.

Alice fixes a positive integer $k$ and chooses a tuple of elements $\bar{a} = (a_1, \ldots, a_k)$. Bob fixes a positive integer $l$ and chooses a tuple of elements $\bar{b} = (b_1, \ldots, b_l)$. These two tuples are public.

**Algorithm.**

- Alice picks a private group word $u = u(x_1, \ldots, x_k)$; then she computes $u_0 = u(a_1, \ldots, a_k)$ and sends the tuple $\bar{b}^{u_0} = (b_1^{u_0}, \ldots, b_l^{u_0})$ to Bob.
- Bob picks a private group word $v = v(y_1, \ldots, y_l)$; then he computes $v_0 = v(b_1, \ldots, b_k)$ and sends the tuple $\bar{a}^{v_0} = (a_1^{v_0}, \ldots, a_k^{v_0})$ to Alice.
- Alice computes

$$u(a_1^{v_0}, \ldots, a_k^{v_0}) u_0^{-1} = u_0^{v_0} u_0^{-1} = [v_0, u_0].$$

- Bob computes
$$v_0 v(b_1^{u_0}, \ldots, b_l^{u_0})^{-1} = v_0 (v_0^{u_0})^{-1} = [v_0, u_0].$$
Now the commutator
$$K = [v_0, u_0]$$
is the secret exchanged key.

**5E.** ***Cryptanalysis of the Anshel–Anshel–Goldfeld protocol.*** The AAG protocol was analyzed by Tsaban in [Tsaban 2015; Ben-Zvi et al. 2018]. We will give his analysis for the reader's convenience because we are going to present an improvement of AAG to make it resistant to such sort of attacks.

The commutator key-exchange protocol uses the Artin braid group $B_n$, $n \in \mathbb{N}$, as its platform group. It was shown in [Tsaban 2015] that the problem of computing the exchanged key reduces, polynomially, to the same problem in matrix groups over finite fields. Now let $G$ be a matrix group and two sets $\{a_1, \ldots, a_k\}$ and $\{b_1, \ldots, b_l\}$ be as in the protocol. Let $A = \mathrm{gp}(a_1, \ldots, a_k)$ and $B = \mathrm{gp}(b_1, \ldots, b_l)$ be subgroups generated by these sets respectively. Also denote by $\mathrm{Alg}(A)$ and $\mathrm{Alg}(B)$ the subalgebras (and so vector spaces) generated by $A$ and $B$ respectively.

The linear span-method by Tsaban works as follows:

(1) Compute bases for the vector spaces of $\mathrm{Alg}(A)$ and $\mathrm{Alg}(B)$.

(2) Solve the following homogeneous system of linear equations in the unknown matrix $x \in \mathrm{Alg}(A)$:
$$b_i \cdot x = x \cdot b_i^{u_0}, \quad i = 1, \ldots, l,$$
a system of linear equations on the coefficients determining the matrix $x$, as a linear combination of the basis of the space $\mathrm{Alg}(A)$.

(3) Fix a basis for the solution space, and pick random solutions until the picked solution $x_0$ is invertible.

(4) Solve the following homogeneous system of linear equations in the unknown matrix $y \in \mathrm{Alg}(B)$:
$$a_j \cdot y = y \cdot a_j^{v_0}, \quad j = 1, \ldots, k,$$
a system of linear equations on the coefficients determining $y$, as a linear combination of the basis of the space $\mathrm{Alg}(A)$.

(5) Fix a basis for the solution space, and pick random solutions until the picked solution $y_0$ is invertible.

(6) Output:
$$[x_0, y_0].$$

It is easy to prove, see [Ben-Zvi et al. 2018], that the output is correct, i.e., $[x_0, y_0] = [u_0, v_0]$. That steps (3) and (5) terminate quickly follows from the invertibility lemma, Lemma 2.2.

**Remark 5.1.** The linear span-method by Tsaban et al. described above is efficiently applicable to schemes based on the intractability of the conjugacy search problem for matrix groups over finite fields. It cannot be directly applied to schemes that use abstract groups or matrices over infinite fields groups as the platforms.

## 6. Marginal subsets

In this section we introduce a new concept that can be effectively used to improve some cryptographic schemes, including algebraic cryptography protocols like AAG and DH. This concept formally generalizes the well-known concept of the marginal subgroup, but it is worth noting that this generalization is very different from the original concept.

The marginal subgroup is determined by the word, and the marginal subset is determined by the word and its chosen value. The set of all marginal subsets is not closed under group-theoretic operations. A marginal subset can be very wild.

Let $F$ be a free group on a countably infinite set $\{x_1, x_2, \dots\}$ and let $W$ be a nonempty subset of $F$. If $w = w(x_1, \dots, x_n) \in W$ and $g_1, \dots, g_n$ are elements of a group $G$, we define the *value* of the word $w$ at $(g_1, \dots, g_n)$ to be $w(g_1, \dots, g_n)$. The subgroup of $G$ generated by all values in $G$ of words in $W$ is called the *verbal subgroup* of $G$ determined by $W$,

$$W(G) = \mathrm{gp}(w(g_1, \dots, g_n) : g_i \in G, w \in W).$$

If $W$ is a nonempty set of words in $x_1, x_2, \dots$ and $G$ is any group, a normal subgroup $N$ is said to be $W$*-marginal* in $G$ if

$$w(g_1, \dots, g_n) = w(u_1 g_1, \dots, u_n g_n)$$

for all $w(x_1, \dots, x_n) \in W$, $g_i \in G$, $u_i \in N$, $1 \le i \le n$. This is equivalent to the requirement $g_i = f_i \pmod N$, $1 \le i \le n$, always implies that $w(g_1, \dots, g_n) = w(f_1, \dots, f_n)$.

In particular, for $n \in \mathbb{N}$, any group word $w = w(x_1, \dots, x_n)$ and any group $G$, a normal subgroup $N$ is said to be $w$-marginal in $G$ if

$$w(g_1, \dots, g_n) = w(u_1 g_1, \dots, u_n g_n)$$

for all $g_i \in G$, $u_i \in N$, $1 \le i \le n$. This is equivalent to the requirement $g_i = f_i \pmod N$, $1 \le i \le n$, always implies that $w(g_1, \dots, g_n) = w(f_1, \dots, f_n)$.

Since every set of $W$-marginal subgroups of $G$ generate a normal subgroup that is also marginal, there is the maximal $W$-marginal (in particular $w$-marginal) subgroup of $G$ denoted by $W^*(G)$ (in particular $w^*(G)$). See [Robinson 1982] for more details about verbal and marginal subgroups.

We introduce a new notion that significantly extends the marginality property. For simplicity we give this notion for the case when $W$ consists of a single word $w$. This notion can be easily extended to any set $W$.

**Definition 6.1.** For $n \in \mathbb{N}$, let $w = w(x_1, \dots, x_n)$ be a group word, $G$ be a group and $\bar{g} = (g_1, \dots, g_n)$ be a tuple of elements of $G$. We say that a tuple $\bar{c} = (c_1, \dots, c_n) \in G^n$ is a *marginal tuple* determined by $w$ and $\bar{g}$ if

$$w(c_1 g_1, \dots, c_n g_n) = w(g_1, \dots, g_n).$$

We will write $\bar{c} \perp w(\bar{g})$ in this case. A set $\overline{C} \subseteq G^n$ is said to be *marginal* with respect to $w$ and $\bar{g}$, and write $\overline{C} \perp w(\bar{g})$, if $\bar{c} \perp w(\bar{g})$ for every tuple $\bar{c} \in \overline{C}$.

**Remark 6.2.** Let $G$ be a group, $w = w(x_1, \dots, x_n)$ be a word and $\bar{g} = (g_1, \dots, g_n)$ be a tuple of elements of $G$. Then the following marginality properties are true for $G$, $w$ and $\bar{g}$:

(1) Each subset of a marginal set is marginal.

(2) The direct power $(w^*)^n$ is marginal.

(3) The component $c_i$ of any marginal tuple $\bar{c}$ can be any element of the group $G$ if $w$ is independent of $x_i$.

(4) The set $C_i,\ i = 1, \dots, n$, consisting of all $i$-th components of all $\bar{c} \in \overline{C}$, is generally not closed with respect to group operations. For example, if $g_i$ occurs in $w(\bar{g})$ all times in the form $g_i^2$ then any element $h \in G$ such that $h^2 = 1$ and $hg_i = g_i h$ can be the $i$-th component of a marginal tuple $\bar{c}$ with trivial other components. But the product of two such elements $h$ cannot be an involution and so this product is out of $C_i$ in the general case.

(5) There are many ways to construct a marginal set. Obviously, we can even construct a nonrecursive marginal set in the case of the infinite group $G$. Below we present a very simple and efficient algorithm for constructing a marginal set using the word $w$.

*A method for constructing the marginal set $\overline{C}, \overline{C} \perp w$, based on $w$.* As we noted in Remark 6.2, the marginal set $\overline{C}, \overline{C} \perp w$, is generally not closed under group operations. This set can be chosen as very wild; for example, it can be computable, but not recursive. We are to develop various methods for creating such sets. We also note that the proposed idea can be established as an improvement of many other cryptographic schemes based on the insolubility of the problem of finding conjugacy in groups to make these schemes resistant to attacks by the linear algebra methods.

Now we give a very simple and efficient algorithm for constructing the marginal set $\overline{C}$ using the word $w$. This method is universal because it does not depend on the structure of $G$.

Let $w = w(a_1, \dots, a_k) = a_1 a_2 \cdots a_k,\ a_i \in G,\ i = 1, \dots, k$, be any expression in the straight form of a fixed element $f \in G$. It is possible that $a_i = a_j$ or $a_i = a_j^{-1}$ for $i \neq j$. Also this expression can be nonreduced. Consider the equation

$$x_1 a_1 x_2 a_2 \cdots x_k a_k = f. \tag{11}$$

Every solution of (11) can be included in a marginal set $\overline{C}, \overline{C} \perp w$. We can fix $i$ and choose any values $x_j = c_j,\ j \neq i,\ c_j \in G$. Then we obtain the solution of (11) by setting

$$x_i = a_{i-1}^{-1} c_{i-1}^{-1} \cdots a_1^{-1} c_1^{-1} f a_k^{-1} c_k^{-1} \cdots a_{i+1}^{-1} c_{i+1}^{-1}. \tag{12}$$

We can also generate a solution of (11) using a sequence of the following random elementary inserts. Suppose we have a solution $(c_1, \dots, c_k)$ of (11). For any $i$ and any random element $d \in G$ we can change $c_i$ to $c_i' = c_i a_i d a_i^{-1}$ and $c_{i+1}$ to $c_{i+1}' = d c_{i+1}$. Then we get a new solution of (11). Continuing this process with random $i$ and $d$, we get a series of new solutions of (11).

**Remark 6.3.** In the case when $G \leq \mathrm{M}(n, \mathbb{F})$ is a matrix group over $\mathbb{F}$, the notion of a marginal set can be naturally generalized to any ring-word (even to any algebra-word). Let $R$ be a free associative algebra on a countably infinite set $\{x_1, x_2, \dots\}$ over a field $\mathbb{F}$, and let $W$ be a nonempty subset of $R$. If $w = w(x_1, \dots, x_n) \in W$ and $u_1, \dots, u_n$ are elements of $M = \mathrm{M}(n, \mathbb{F})$, we define the *value* of the word $w$ at $(u_1, \dots, u_n)$ to be $w(u_1, \dots, u_n)$. Let $\bar{g} = (g_1, \dots, g_n)$ be a tuple of elements of $G$. We say that a tuple $\bar{c} = (c_1, \dots, c_n) \in M^n$ is a *marginal tuple* determined by $w$ and $\bar{g}$ if

$$w(c_1 g_1, \dots, c_n g_n) = w(g_1, \dots, g_n).$$

Other generalizations when we use the ring $T$ instead of the group $G$ or use more general operations instead of multiplication on the left side are also possible.

## 7. Improved versions of the AAG and Ko et al. cryptographic protocols

Suppose two correspondents Alice and Bob want to exchange a key. They agree about a group $G$ given by a finite set of generators that is used as the platform. It is supposed that $G$ is equipped with an efficient normal form of its elements and the main group operations can be computed efficiently. All the information about $G$, the normal form and efficient algorithms to compute products of elements, their inversions and normal forms, is public. In particular, the word problem is efficiently solvable for $G$.

**7A.** ***An improved version of the AAG key exchange protocol.*** To exchange a key the correspondents act as follows.

Alice fixes a positive integer $k$ and chooses a tuple of elements $\bar{a} = (a_1, \dots, a_k)$. Then she picks up a private group word $u = u(x_1, \dots, x_k)$ and computes $u(\bar{a}) = u(a_1, \dots, a_k)$. Also she finds a marginal set $\overline{C} \subseteq G^k$, $\overline{C} \perp u(\bar{a})$.

Bob fixes a positive integer $l$ and chooses a tuple of elements $\bar{b} = (b_1, \dots, b_l)$. Then he picks up a private group word $v = v(y_1, \dots, y_l)$ and computes $v(\bar{b}) = v(b_1, \dots, b_l)$. Also he finds a marginal set $\overline{D} \subseteq G^l$, $\overline{D} \perp v(\bar{b})$.

Alice publishes elements $a_1, \dots, a_k$ as $a_{\pi(1)}, \dots, a_{\pi(k)}$, where $\pi \in \mathbb{S}_k$ is a random permutation. The same permutation is applied to the corresponding tuples $\bar{c} \in \overline{C}$.

Bob acts in the similar way.

*Virtual and hidden elements.* Alice can also introduce a virtual element $h$ that is not used in the expression for $u(\bar{a})$. Then she add a new random component to any $\bar{c} \in \overline{C}$, $\overline{C} \perp w$. She can add many such components with aim to hide the length of the word $u$, or to hide equality (12), or choose some element $h$ with huge centralizer as well as with small centralizer, to make solution of the problem more difficult for an intruder. Bob acts similarly.

Also Alice can hide some elements $a_i$ as follows. Let $a_i = a_j$ and the corresponding components $c_i = c_j$ for all $\bar{c} \in \overline{C}$. Then Alice does not publish $a_j$ and removes the $j$-component from every $\bar{c}$. Bob acts similarly.

These two operations are recommended. After these operations the parameters $k$ and $l$ can be changed to $k'$ and $l'$ respectively.

Alice publishes elements $a_1, \dots, a_{k'}$ as $a_{\pi(1)}, \dots, a_{\pi(k')}$, where $\pi \in \mathbb{S}_{k'}$ is a random permutation. The same permutation is applied to the corresponding tuples $\bar{c} \in \overline{C}$.

Bob acts in the similar way.

Alice publishes elements $a_1, \dots, a_{k'}$ as $a_{\pi(1)}, \dots, a_{\pi(k')}$, where $\pi \in \mathbb{S}_{k'}$ is a random permutation. The same permutation is applied to the corresponding tuples $\bar{c} \in \overline{C}$, and they are published.

Bob acts in the similar way.

**Algorithm.**

- Alice picks a private tuple $\bar{d} = (d_1, \dots, d_{l'}) \in \overline{D}$ and computes $\overline{db} = (d_1 b_1, \dots, d_{l'} b_{l'})$. Then she sends the tuple $\overline{db}^{u(\bar{a})} = ((d_1 b_1)^{u(\bar{a})}, \dots, (d_{l'} b_{l'})^{u(\bar{a})})$ to Bob.
- Bob picks a private tuple $\bar{c} = (c_1, \dots, c_{k'}) \in \overline{C}$ and computes $\overline{ca} = (c_1 a_1, \dots, c_{k'} a_{k'})$. Then he sends the tuple $\overline{ca}^{v(\bar{b})} = ((c_1 a_1)^{v(\bar{b})}, \dots, (c_{k'} a_{k'})^{v(\bar{b})})$ to Alice.
- Alice computes

$$u((c_1 a_1)^{v(\bar{b})}, \dots, (c_k a_k)^{v(\bar{b})}) = u(\bar{a})^{-1} u(c_1 a_1, \dots, c_k a_k)^{v(\bar{b})} = [u(\bar{a}), v(\bar{b})].$$

- Bob computes similarly

$$v((d_1b_1)^{u(\bar a)},\dots,(d_lb_l)^{u(\bar a)})^{-1}v(\bar b)=(v(d_1b_1,\dots,d_lb_l)^{u(\bar a)})^{-1}v(\bar b)=[u(\bar a),v(\bar b)].$$

  Now the commutator

$$K=[u(\bar a),v(\bar b)]$$

  is the secret exchanged key.

**Definition 7.1.** The conjugacy-membership problem is solvable for $G$ with respect to $\overline{C}\subseteq^k$ if there is an algorithm that decides for any two tuples $\bar a=(a_1,\dots,a_k)$ and $\bar f=(f_1,\dots,f_k)$ of elements of $G$ whether or not there exists an element $y\in G$ such that $(f_1^y a_1^{-1},\dots,f_k^y a_k^{-1})\in\overline{C}$. In short, is there an element $y\in G$ such that $\bar f^y\bar a^{-1}\in\overline{C}$? The corresponding problem, which is a mixture of conjugacy and membership problems, is the question of the existence of an algorithm that finds a solution, if such a solution exists.

The proposed version of the AAG protocol is based on intractability of the mixed conjugacy-membership search problem when $\overline{C}$ is a marginal set, $\overline{C}\perp u(a_1,\dots,a_k)$, for the unknown word $u(x_1,\dots,x_n)$ (or similarly when $\overline{D}$ is a marginal set, $\overline{D}\perp v(b_1,\dots,b_l)$). Indeed, suppose that an intruder finds $\bar c'\in\overline{C}$ and $y\in G$ such that $\overline{ca}^{v(\bar b)}=\overline{c'a}^y$, and similarly he finds $\bar d'\in\overline{D}$ and $x\in G$ such that $\overline{db}^{u(\bar a)}=\overline{d'b}^x$. Then $[x,y]=[u(\bar a),v(\bar b)]$ as in the original version.

There are other problems that should probably be addressed first. The presence of virtual and hidden elements does not allow us to calculate the lengths of $u$ and $v$. We also note that each solution of (11) is also a solution to each equation of the form $a_ia_{i+1}\cdots a_ka_1\cdots a_{i-1}=f$, $i=2,\dots,k$, and possibly some other equations. Therefore, the open data does not allow us to unambiguously restore $f^{v(\bar b)}$, even if the attacker knows the length of $v$ and all the letters $v(\bar b)$ with their multiplicity.

**7B.** ***An improved version of the Ko et al. key exchange protocol.*** To exchange a key the correspondents act as follows.

Let $G$ be a group. Alice and Bob agree about a public element $g\in G$. Let $A$ and $B$ be two finitely generated elementwise commuting subgroups of $G$. This data is public.

Alice fixes a positive integer $k$ and chooses a tuple of elements $\bar f=(f_1,\dots,f_k)$ such that $g\in\mathrm{gp}(f_1,\dots,f_k)$. Then she picks a private group word $u=u(x_1,\dots,x_k)$ such that $g=u(\bar f)$. Also she finds a marginal set $\overline{C}\subseteq G^k$, $\overline{C}\perp u(\bar f)$. Alice publishes $\overline{C}$.

Bob fixes a positive integer $l$ and chooses a tuple of elements $\bar f'=(f_1',\dots,f_l')$ such that $g\in\mathrm{gp}(f_1',\dots,f_l')$. Then he picks a private group word $v=v(x_1,\dots,x_l)$ such that $g=v(\bar f')$. Also he finds a marginal set $\overline{D}\subseteq G^l$, $\overline{D}\perp v(\bar f')$. Bob publishes $\overline{D}$.

If $G$ is a matrix group, the words $u$ and $v$ can be ring-words.

**Algorithm.**

- Alice chooses a private tuple $\bar h=(h_1,\dots,h_k)\in\mathrm{C}_G(B)^k$ and computes $\tilde f=(f_1h_1,\dots,f_kh_k)$. Then she publishes $\tilde f$.
- Bob chooses a private tuple $\bar h'=(h_1',\dots,h_l')\in\mathrm{C}_G(A)^l$ and computes $\tilde f'=(f_1'h_1',\dots,f_l'h_l')$. Then he publishes $\tilde f'$.
- Alice picks a random tuple $\bar d=(d_1,\dots,d_l)\in\overline{D}$ and computes $\bar d\,\tilde f'=(d_1\tilde f_1',\dots,d_l\tilde f_l')$. She also chooses a random private element $a\in A$. Then she sends $(\bar d\,\tilde f')^a=((d_1\tilde f_1')^a,\dots,(d_l\tilde f_l')^a)$ to Bob.

- Bob picks a random tuple $\bar{c} = (c_1, \dots, c_k) \in \overline{C}$ and computes $\bar{c}\tilde{f} = (c_1\tilde{f}_1, \dots, c_k\tilde{f}_k)$. He chooses a random private element $b \in B$. Then he sends $(\bar{c}\tilde{f})^b = ((c_1\tilde{f}_1)^b, \dots, (c_k\tilde{f}_k)^b)$ to Alice.
- Alice computes

$$(\bar{c}\tilde{f})^b\bar{h}^{-1} = ((c_1\tilde{f}_1)^b h_1^{-1}, \dots, (c_k\tilde{f}_k)^b h_k^{-1}) = ((c_1 f_1)^b, \dots, (c_k f_k)^b) = (\bar{c}\bar{f})^b.$$

- Alice computes

$$u((\bar{c}\bar{f})^b) = u(\bar{c}\bar{f})^b = u(\bar{f})^b = g^b.$$

- Bob computes

$$(\bar{d}\tilde{f}')^a(\bar{h}')^{-1} = ((d_1\tilde{f}_1')^a(h_1')^{-1}, \dots, (d_l\tilde{f}_l')^a(h_l')^{-1}) = ((d_1 f_1')^a, \dots, (d_l f_l')^a) = (\bar{d}\bar{f}')^a.$$

- Bob computes

$$v((\bar{d}\bar{f}')^a) = v(\bar{d}\bar{f}')^a = v(\bar{f}')^a = g^a.$$

- Alice computes $K_A = (g^b)^a = g^{ab}$.
- Bob computes $K_B = (g^a)^b = g^{ab}$, and

$$K = K_A = K_B = g^{ab}$$

is the secret exchanged key.

**Remark 7.2.** Alice publishes instead of $f_1, \dots, f_k$ changed elements $\tilde{f}_1, \dots, \tilde{f}_k$. This is done in order to make it difficult for a potential cracker to select the expression $u(f_1, \dots, f_k)$. Since each element $h_i$ lies in $\mathrm{C}_G(B)$, the element $b \in B$ acts on $h_i$ trivially. Alice may exclude $h_i^b = h_i$ from $c_i\tilde{f}^b$ and get $c_i f_i^b$. Some of the elements $f_1, \dots, f_k$ are virtual. This means that the value $u(f_1, \dots, f_k)$ does not depend on them. Therefore, the choice in the marginal set $\overline{C}$ of the corresponding components can be carried out randomly. It is also possible that for $i \neq j$ we have $f_i = f_j$. Then both of these elements are published, and the corresponding elements $h_i, c_i$ and $h_j, c_j$ are chosen independently. If an element $f_i$ occurs several times in the expression $u(f_1, \dots, f_k)$, then it is published once. The elements $h_i$ and $c_i$ corresponding to it are also selected once.

All of the above also holds true for Bob to select parameters.

We show a toy example of the just-considered improved version of the key exchange protocol with simple parameters.

**Example 7.3.** First we will give a symbolic description of the protocol.

Let $G = \mathrm{GL}(6, \mathbb{Z})$, and let $A, B \leq G$ be two elementwise permutable subgroups of $G$ given by their generating sets $\{a_1, \dots, a_k\}$ and $\{b_1, \dots, b_l\}$ respectively, and $g \in G$. This data is public. Suppose that

$$u(x_1, x_2, x_3) = [x_1, x_2]x_2^{-1} + x_2x_3 - x_1$$

is a ring word in which the variables $x_1$ and $x_2$ take invertible values.

We choose a pair of elements $f_1$ and $f_2$ of $G$ so that the element

$$f_3 = f_2^{-1}g - f_2^{-1}[f_1, f_2]f_2^{-1} + f_2^{-1}f_1$$

is invertible.

Then

$$g = u(f_1, f_2, f_3) = [f_1, f_2]f_2^{-1} + f_2 f_3 - f_1.$$

For $c_1, c_2, c_3 \in \mathrm{M}(6, \mathbb{Z})$, the equality

$$u(c_1 f_1, c_2 f_2, c_3 f_3) = u(f_1, f_2, f_3)$$

is true if and only if

$$c_3 = f_2^{-1} c_2^{-1} (g - [c_1 f_1, c_2 f_2] f_2^{-1} c_2^{-1} + c_1 f_1) f_3^{-1}. \tag{13}$$

The formula (13) describes the full marginal set $\widetilde{C} \perp u(f_1, f_2, f_3)$. Then Alice constructs an infinite marginal set $\overline{C}_3 = \{(c_1(i), c_2(i), c_3(i)) : i = 1, 2, \dots\}$, choosing the elements $c_1(i)$ and $c_2(i)$ in $G$ and calculating $c_3(i)$ according to (13).

Then Alice randomly chooses the elements $h_1, h_2, h_3 \in \mathrm{C}_G(B)$ (where $\mathrm{C}_G(B)$ denotes the centralizer of $B$ in $G$) and calculates the elements $\tilde{f}_i = f_i h_i$ for $i = 1, 2, 3$. She also chooses a number $k \geq 3$ and the random virtual elements $\tilde{f}_4, \dots, \tilde{f}_k \in G$. For each $i = 4, 5, \dots$, she takes the random elements $c_4(i), \dots, c_k(i) \in G$ and publishes the constructed marginal set $\overline{C} = \{(c_1(i), c_2(i), c_3(i), c_4(i), \dots, c_k(i)) : i = 1, 2, \dots\}$. In practice, she also applies a random permutation to the indices of the tuple $(\tilde{f}_1, \dots, \tilde{f}_k)$ and to each of the corresponding tuples from $\overline{C}$, so as not to show which ones are virtual. To simplify the recording, we do not do this hereinafter.

In continuation of the algorithm Bob picks a random element $b \in B$, chooses randomly $\bar{c}(q) \in \overline{C}$, calculates and publishes the elements

$$(c_i(q)\tilde{f}_i)^b \quad \text{for } i = 1, \dots, k.$$

Alice calculates

$$(c_i(q)\tilde{f}_i)^b h_i^{-1} = (c_i(q) f_i)^b \quad \text{for } i = 1, 2, 3. \tag{14}$$

Then she obtains

$$u((c_1(q) f_1)^b, (c_2(q) f_2)^b, (c_3(q) f_3)^b) = u(c_1(q) f_1, c_2(q) f_2, c_3(q) f_3)^b = u(f_1, f_2, f_3)^b = g^b. \tag{15}$$

Also Alice randomly chooses an element $a \in A$ and computes the key: $(g^b)^a = g^{ab}$.

Bob acts the same way.

Next, we will give numerical values for the protocol parameters in our example. We set

$$A = \mathrm{gp}(t_{13}, t_{31}, t_{35}, t_{53}), \quad B = \mathrm{gp}(t_{24}, t_{42}, t_{46}, t_{64}),$$

where $t_{ij} = e + e_{ij}$, $i \neq j$, is a transvection and, for each pair $ij$, $e_{ij}$ is an elementary matrix that differs from the zero matrix by one element 1 that stands in the $ij$-position. Obviously, $A$ and $B$ are elementwise permutable. We also set

$$g = e + e_{12} + e_{23} + e_{34} + e_{45} + e_{56}, \quad f_1 = t_{23}, \quad f_2 = t_{34}.$$

Then

$$f_3 = e + e_{12} + 2e_{23} + e_{34} + e_{45} + e_{56} - e_{24} - e_{35},$$

$$f_3^{-1} = e - e_{12} - 2e_{23} - e_{34} - e_{45} - e_{56} + 2e_{13} + 3e_{24} + 2e_{35} + e_{46} - 3e_{14} - 5e_{25} - 2e_{36} + 5e_{15} + 5e_{26} - 5e_{16}.$$

Alice picks

$$h_1 = t_{31}, \quad h_2 = t_{13} t_{53}^{-1}, \quad h_3 = t_{15}^2 t_{53}.$$

Then

$$\tilde{f}_1 = f_1 h_1 = e + e_{23} + e_{21} + e_{31},$$
$$\tilde{f}_2 = f_2 h_2 = e + e_{34} + e_{13} - e_{53},$$
$$\tilde{f}_3 = f_3 h_3 = e - e_{33} + e_{12} + 2e_{23} + e_{34} + e_{45} + e_{56} - 2e_{13} - e_{24} - e_{35} + 2e_{15} + e_{43} + e_{53}.$$

Bob chooses randomly $\bar{c}(q_0) \in \overline{C}$. For example he takes

$$c_1(q_0) = e - e_{34} + e_{21},$$
$$c_2(q_0) = t_{23},$$
$$c_3(q_0) = e - 2e_{22} + 6e_{23} - e_{34} + e_{35} - 10e_{24} + 16e_{25} - e_{36} - 16e_{26} + 2e_{21}, ,$$
$$c_i(q_0) \quad \text{for } i = 4, \ldots, k$$

(we do not specify these virtual elements).

Then Bob picks $b = t_{24} \in B$, he calculates the elements

$$(c_1(q_0)\tilde{f}_1)^b = e + e_{23} - e_{34} + 2e_{21} + e_{31},$$
$$(c_2(q_0)\tilde{f}_2)^b = e + e_{34} + e_{13} - e_{53},$$
$$(c_3(q_0)\tilde{f}_3)^b = e + e_{12} + e_{23} + e_{45} + e_{56} - 2e_{13} - 3e_{24} - e_{35} + 5e_{25} - e_{14} + 2e_{15} + 2e_{21} + e_{43} + e_{53},$$
$$(c_i(q_0)\tilde{f}_i)^b \quad \text{for } i = 4, \ldots, k.$$

Then he publishes

$$(c_1(q_0)\tilde{f}_1)^b, \ldots, (c_k(q_0)\tilde{f}_k)^b).$$

Suppose that Alice picks $a = t_{35} \in A$.

Alice calculates

$$(c_1(q_0) f_1)^b = (c_1(q_0)\tilde{f}_1)^b h_1^{-1} = e + e_{23} - e_{34} + e_{21},$$
$$(c_2(q_0) f_2)^b = (c_2(q_0)\tilde{f}_2)^b h_2^{-1} = e + e_{34},$$
$$(c_3(q_0) f_3)^b = (c_3(q_0)\tilde{f}_3)^b h_3^{-1} = e + e_{12} + 4e_{23} + e_{45} + e_{56} - 3e_{24} - e_{35} - e_{14} + e_{25} + 2e_{21}.$$

By (15) she obtains that

$$g^b = e + e_{12} + e_{23} + e_{34} + e_{45} + e_{56} - e_{14} + e_{25}.$$

Then

$$(g^b)^a = g^{ab} = e + e_{12} + e_{23} + e_{34} + e_{45} + e_{56} + e_{36} - e_{14} \tag{16}$$

is the exchanged key.

Bob takes a ring word

$$v(x_1, x_2, x_3, x_4)) = x_1 x_2 - x_3 + x_4$$

and elements

$$f_1' = e + e_{24} - e_{21}, \quad f_3' = e - e_{23} - e_{45} - e_{56} + e_{24} - e_{21} - e_{32},$$
$$f_2' = e + e_{34}, \qquad f_4' = e + e_{12} - e_{32}.$$

Then

$$g = v(f'_1, \dots, f'_4) = f'_1 f'_2 - f'_3 + f'_4.$$

Bob picks the following random elements in $\mathrm{C}_G(A)$:

$$h'_1 = e + e_{24}, \quad h'_2 = e - e_{42}, \quad h'_3 = e + e_{46} + e_{42}, \quad h'_4 = e + 2e_{24}.$$

Then he computes

$$\begin{aligned}
\tilde{f}'_1 &= f'_1 h'_1 = e - e_{21} + 2e_{24},\\
\tilde{f}'_2 &= f'_2 h'_2 = e + e_{34} - e_{42} - e_{32},\\
\tilde{f}'_3 &= f'_3 h'_3 = e - e_{23} - e_{45} - e_{56} + e_{24} - e_{21} - e_{32} + e_{46} + e_{42} + e_{26} + e_{22},\\
\tilde{f}'_4 &= f'_4 h'_4 = e + e_{12} - e_{32} + 2e_{24} + 2e_{14} - 2e_{34}.
\end{aligned}$$

The full marginal set $\widetilde{D} \perp v(f'_1, \dots, f'_4)$ is described by

$$d_4 = (g - d_1 f'_1 d_2 f'_2 + d_3 f'_3)(f'_4)^{-1}.$$

Then one has

$$v(d_1 f'_1, \dots, d_4 f'_4) = v(f'_1, \dots, f'_4).$$

Bob constructs an infinite marginal set $\overline{D}_4 = \{(d_1(i), \dots, d_4(i)) : i = 1, 2, \dots\}$.

Bob chooses a number $l \geq 4$ and the random virtual elements $f'_5, \dots, f'_l$. For each $i = 5, 6, \dots$ he takes the random elements $d_5(i), \dots, d_l(i)$ and publishes the constructed marginal set

$$\overline{D} = \{d_1(i), \dots, d_4(i), d_5(i), \dots, d_l(i) : i = 1, 2, \dots\}.$$

In practice, he also applies a random permutation to the indices of the tuple $(\tilde{f}'_1, \dots, \tilde{f}'_l)$ and to each of the corresponding tuples from $\overline{D}$, so as not to show which ones are virtual. To simplify the recording, we do not do this hereinafter.

Alice chooses $\bar{d}(p_0) \in \overline{D}$:

$$\begin{aligned}
&d_1(p_0) = e + e_{32}, d_2(p_0) = e - e_{23}, d_3(p_0) = e - e_{45} + e_{13},\\
&d_4(p_0) = e + e_{22} + e_{33} + e_{23} - e_{45} + e_{13} + e_{24} + e_{46} - e_{32} + e_{31}, d_i(p_0)
\end{aligned}$$

for $i = 5, \dots, l$. She computes

$$\begin{aligned}
(d_1(p_0)\tilde{f}'_1)^a &= e - 2e_{34} + 2e_{24} - e_{21} + e_{32} - e_{31},\\
(d_2(p_0)\tilde{f}'_2)^a &= e + e_{22} - e_{23} + e_{34} - e_{24} + e_{25} - e_{32} - e_{42},\\
(d_3(p_0)\tilde{f}'_3)^a &= e + e_{22} - e_{23} - e_{45} - e_{56} + e_{24} + e_{46} + e_{25} - e_{36} + e_{26} - e_{21} - e_{32} + e_{42},\\
(d_4(p_0)\tilde{f}'_4)^a &= e + e_{22} + e_{33} + e_{23} - e_{45} + e_{13} + e_{24} - e_{35} + e_{46} - e_{25} - e_{15} - e_{32} + e_{31},
\end{aligned}$$

and $(d_i(p_0)\tilde{f}'_i)^a$ for $i = 5, \dots, l$.

Then she publishes

$$((d_1(p_0)\tilde{f}'_1)^a, \dots, (d_l(p_0)\tilde{f}'_l)^a).$$

Bob computes

$$(d_1(p_0)f_1')^a = (d_1(p_0)\tilde{f}_1')^a(h_1')^{-1} = e+e_{34}+e_{24}-e_{21}+e_{32}-e_{31},$$
$$(d_2(p_0)f_2')^a = (d_2(p_0)\tilde{f}_2')^a(h_2')^{-1} = e-e_{23}+e_{34}-e_{24}+e_{25},$$
$$(d_3(p_0)f_3')^a = (d_3(p_0)\tilde{f}_3')^a(h_3')^{-1} = e-e_{12}-e_{23}-2e_{45}-e_{56}+e_{13}+e_{24}+e_{46}+e_{25}-e_{36}-e_{15}-e_{21}-e_{32},$$
$$(d_4(p_0)f_4')^a = (d_4(p_0)\tilde{f}_4')^a(h_4')^{-1} = e+e_{33}+e_{23}-e_{45}+e_{13}+e_{24}+e_{46}-e_{35}-e_{25}-e_{15}-2e_{32}+e_{31}.$$

Now he obtains

$$v((d_1(p_0)f_1')^a, \dots, (d_4(p_0)f_4')^a) = v(d_1(p_0)f_1', \dots, d_4(p_0)f_4')^a = g^a, \tag{17}$$

and computes $(g^a)^b = g^{ab}$; see (16).

## References


[Anshel et al. 1999] I. Anshel, M. Anshel, and D. Goldfeld, "An algebraic method for public-key cryptography", *Math. Res. Lett.* **6**:3-4 (1999), 287–291. MR Zbl

[Baba et al. 2011] S. Baba, S. Kotyada, and R. Teja, "A non-Abelian factorization problem and an associated cryptosystem", report 2011/048, Cryptology ePrint Archive, 2011, available at https://eprint.iacr.org/2011/048.pdf.

[Ben-Zvi et al. 2018] A. Ben-Zvi, A. Kalka, and B. Tsaban, "Cryptanalysis via algebraic spans", pp. 255–274 in *Advances in cryptology—CRYPTO 2018, Part I*, edited by H. Shacham and A. Boldyreva, Lecture Notes in Comput. Sci. **10991**, Springer, 2018. MR Zbl

[Bhunia et al. 2019] S. Bhunia, A. Mahalanobis, P. Shinde, and A. Singh, "The MOR cryptosystem in classical groups with a Gaussian elimination algorithm for symplectic and orthogonal groups", in *Modern cryptography*, edited by M. Domb, IntechOpen, 2019.

[Bigelow 2001] S. J. Bigelow, "Braid groups are linear", *J. Amer. Math. Soc.* **14**:2 (2001), 471–486. MR Zbl

[Cheon and Jun 2003] J. H. Cheon and B. Jun, "A polynomial time algorithm for the braid Diffie–Hellman conjugacy problem", pp. 212–225 in *Advances in cryptology—CRYPTO 2003*, edited by D. Boneh, Lecture Notes in Comput. Sci. **2729**, Springer, 2003. MR Zbl

[Fine et al. 2016] B. Fine, A. I. S. Moldenhauer, and G. Rosenberger, "Cryptographic protocols based on Nielsen transformations", *J. Comput. and Comm.* **4**:12 (2016), 63–107.

[Garber et al. 2006] D. Garber, S. Kaplan, M. Teicher, B. Tsaban, and U. Vishne, "Length-based conjugacy search in the braid group", pp. 75–87 in *Algebraic methods in cryptography*, edited by L. Gerritzen et al., Contemp. Math. **418**, Amer. Math. Soc., Providence, RI, 2006. MR Zbl

[Hofheinz and Steinwandt 2002] D. Hofheinz and R. Steinwandt, "A practical attack on some braid group based cryptographic primitives", pp. 187–198 in *Public key cryptography—PKC 2003*, edited by Y. G. Desmedt, Lecture Notes in Comput. Sci. **2567**, Springer, 2002. MR

[Hughes 2002] J. Hughes, "A linear algebraic attack on the AAFG1 braid group cryptosystem", pp. 176–189 in *ACISP 2002: Information security and privacy* (Melbourne, 2002), edited by L. Batten and J. Seberry, Lecture Notes in Computer Science **2384**, Springer, 2002. Zbl

[Hughes and Tannenbaum 2002] J. Hughes and A. Tannenbaum, "Length-based attacks for certain group based encryption rewriting systems", report 2003/102, Cryptology ePrint Archive, 2002, available at https://eprint.iacr.org/2003/102.pdf.

[Kahrobaei and Khan 2006] D. Kahrobaei and B. Khan, "NIS05-6: a non-commutative generalization of ElGamal key exchange using polycyclic groups", pp. 1–5 in *IEEE Globecom 2006*, 2006.

[Ko et al. 2000] K. H. Ko, S. J. Lee, J. H. Cheon, J. W. Han, J.-s. Kang, and C. Park, "New public-key cryptosystem using braid groups", pp. 166–183 in *Advances in cryptology—CRYPTO 2000* (Santa Barbara, CA 2000), edited by M. Bellare, Lecture Notes in Comput. Sci. **1880**, Springer, 2000. MR Zbl

[Krammer 2002] D. Krammer, "Braid groups are linear", *Ann. of Math.* (2) **155**:1 (2002), 131–156. MR Zbl

[Mahalanobis 2008] A. Mahalanobis, "A simple generalization of the ElGamal cryptosystem to non-abelian groups", *Comm. Algebra* **36**:10 (2008), 3878–3889. MR Zbl

[Mahalanobis 2012] A. Mahalanobis, "A simple generalization of the ElGamal cryptosystem to non-abelian groups II", *Comm. Algebra* **40**:9 (2012), 3583–3596. MR Zbl

[Mahalanobis 2015] A. Mahalanobis, "The MOR cryptosystem and extra-special $p$-groups", *J. Discrete Math. Sci. Cryptogr.* **18**:3 (2015), 201–208. MR Zbl

[Menezes and Vanstone 1992] A. J. Menezes and S. A. Vanstone, "A note on cyclic groups, finite fields, and the discrete logarithm problem", *Appl. Algebra Engrg. Comm. Comput.* **3**:1 (1992), 67–74. MR Zbl

[Menezes and Wu 1997] A. J. Menezes and Y.-H. Wu, "The discrete logarithm problem in GL($n$, $q$)", *Ars Combin.* **47** (1997), 23–32. MR Zbl

[Monico 2016] C. Monico, "Cryptanalysis of a matrix-based MOR system", *Comm. Algebra* **44**:1 (2016), 218–227. MR Zbl

[Myasnikov and Roman'kov 2015] A. Myasnikov and V. Roman'kov, "A linear decomposition attack", *Groups Complex. Cryptol.* **7**:1 (2015), 81–94. MR Zbl

[Myasnikov and Ushakov 2007] A. D. Myasnikov and A. Ushakov, "Length based attack and braid groups: cryptanalysis of Anshel–Anshel–Goldfeld key exchange protocol", pp. 76–88 in *Public key cryptography—PKC 2007*, edited by T. Okamoto and X. Wang, Lecture Notes in Comput. Sci. **4450**, Springer, 2007. MR Zbl

[Myasnikov et al. 2005] A. Myasnikov, V. Shpilrain, and A. Ushakov, "A practical attack on a braid group based cryptographic protocol", pp. 86–96 in *Advances in cryptology—CRYPTO 2005*, edited by V. Shoup, Lecture Notes in Comput. Sci. **3621**, Springer, 2005. MR

[Myasnikov et al. 2006] A. Myasnikov, V. Shpilrain, and A. Ushakov, "Random subgroups of braid groups: an approach to cryptanalysis of a braid group based cryptographic protocol", pp. 302–314 in *Public key cryptography—PKC 2006*, edited by M. Yung et al., Lecture Notes in Comput. Sci. **3958**, Springer, 2006. MR

[Myasnikov et al. 2008] A. Myasnikov, V. Shpilrain, and A. Ushakov, *Group-based cryptography*, Birkhäuser, Basel, 2008. MR

[Myasnikov et al. 2011] A. Myasnikov, V. Shpilrain, and A. Ushakov, *Non-commutative cryptography and complexity of group-theoretic problems*, Mathematical Surveys and Monographs **177**, Amer. Math. Soc., Providence, RI, 2011. MR

[Paeng et al. 2001] S.-H. Paeng, K.-C. Ha, J. H. Kim, S. Chee, and C. Park, "New public key cryptosystem using finite nonabelian groups", pp. 470–485 in *Advances in cryptology—CRYPTO 2001* (Santa Barbara, CA, 2001), edited by J. Kilian, Lecture Notes in Comput. Sci. **2139**, Springer, 2001. MR Zbl

[Peker 2014] Y. K. Peker, "A new key agreement scheme based on the triple decomposition problem", *Int. J. Netw. Secur.* **16**:6 (2014), 426–4360.

[Robinson 1982] D. J. S. Robinson, *A course in the theory of groups*, Graduate Texts in Mathematics **80**, Springer, 1982. MR Zbl

[Roman'kov 2012] V. A. Roman'kov, *Introduction to cryptography*, Forum, Moscow, 2012. In Russian.

[Roman'kov 2013a] V. A. Roman'kov, *Algebraic cryptography*, Omsk State University, 2013. In Russian.

[Roman'kov 2013b] V. A. Roman'kov, "Cryptanalysis of some schemes applying automorphisms", *Prikl. Diskr. Mat.* **21** (2013), 35–51. In Russian.

[Roman'kov 2016] V. Roman'kov, "A nonlinear decomposition attack", *Groups Complex. Cryptol.* **8**:2 (2016), 197–207. MR

[Roman'kov 2018a] V. Roman'kov, "Two general schemes of algebraic cryptography", *Groups Complex. Cryptol.* **10**:2 (2018), 83–98. MR

[Roman'kov 2018b] V. A. Roman'kov, *Essays in algebra and cryptology: algebraic cryptanalysis*, Omsk State University, 2018. In Russian.

[Roman'kov 2019a] V. Roman'kov, "An improved version of the AAG cryptographic protocol", *Groups Complex. Cryptol.* **11**:1 (2019), 35–41. MR

[Roman'kov 2019b] V. A. Roman'kov, "Discrete logarithm for nilpotent groups and cryptanalysis of polylinear cryptographic system", *Prikl. Diskr. Mat. Suppl.* **12** (2019), 154–160. In Russian.

[Roman'kov 2019c] V. A. Roman'kov, "Efficient methods of algebraic cryptanalysis and protection against them", *Prikl. Diskr. Mat. Suppl.* **12** (2019), 154–160. In Russian.

[Roman'kov 2019d] V. A. Roman'kov, "Linear algebra methods in cryptanalysis and protection against them", *Herald of Omsk University* **24**:3 (2019), 21–30. In Russian.

[Roman'kov and Obzor 2018] V. A. Roman'kov and A. A. Obzor, "A nonlinear decomposition method in analysis of some encryption schemes using group automorphisms", *Prikl. Diskr. Mat.* **41** (2018), 38–45. In Russian.

[Romsy 2011] M. Romsy, "Adaption of Pollard's kangaroo algorithm to the FACTOR problem", report 2011/483, Cryptology ePrint Archive, 2011, available at https://eprint.iacr.org/2011/483.pdf.

[Sidelnikov et al. 1993] V. M. Sidelnikov, M. A. Cherepnev, and V. V. Yashchenko, "Public key distribution systems based on noncommutative semigroups", *Dokl. Akad. Nauk* **332**:5 (1993), 566–567. In Russian. MR

[Stanek 2011] M. Stanek, "Extending baby-step giant-step algorithm for FACTOR problem", report 2011/059, Cryptology ePrint Archive, 2011, available at https://eprint.iacr.org/2011/059.pdf.

[Tsaban 2015] B. Tsaban, "Polynomial-time solutions of computational problems in noncommutative-algebraic cryptography", *J. Cryptology* **28**:3 (2015), 601–622. MR Zbl

Received 9 Nov 2019. Revised 2 Mar 2020.

VITALIĬ ROMAN'KOV:

romankov48@mail.ru

Mathematical Center, Sobolev Institute of Mathematics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

msp

Moscow Journal of Combinatorics and Number Theory
Vol. 9, No. 2, 2020
https://doi.org/10.2140/moscow.2020.9.147

# On the behavior of power series with positive completely multiplicative coefficients

Oleg A. Petrushov

We consider power series with positive completely multiplicative coefficients. We obtain a large family of power series that have the unit circle as natural boundary, as well as new Ω-theorems for power series with positive completely multiplicative coefficients when the argument tends to roots of unity. Also Ω-estimates for some partial sums of completely multiplicative functions are given.

## 1. Introduction

In this paper we study power series with completely multiplicative coefficients. Power series with coefficients that have some arithmetical structure possess interesting properties. Most of the power series with arithmetical coefficients converge on the unit disc but have no continuation beyond the unit circle. Moreover they usually have interesting properties when $z$ tends to the unit circle along a radius.

The results of [Petrushov 2014; 2015a] are about two specific and important series $\sum_n \mu(n)z^n$ and $\sum_n \mu^2(n)z^n$. For example in [Petrushov 2015a] we exposed a connection between the asymptotic behavior of the series $\sum_{n>0} \mu^2(n)z^n$ as $z$ tends to $e^{2\pi i\beta}$ along its radius and the Diophantine properties of $\beta$, namely its irrationality exponent when $\beta$ is irrational and its denominator when $\beta$ is rational. A similar study for the series $\sum_{n>0} \mu(n)z^n$ was performed in [Petrushov 2014] with less-striking conclusions.

The specific problem of determining the analytic behavior of power series with multiplicative coefficients was posed by W. Schwarz in the Oberwolfach Meeting on Number Theory. L. G. Lucht [1981] proved that for an extensive set of multiplicative functions $\alpha(n)$ the unit circle is the natural boundary of the series $\sum_{n=1}^{\infty} \alpha(n)z^n$. The set is defined by some complicated conditions. In particular it requires the existence of a complex number $s$ with nonnegative real part, a slowly oscillating function $l(x)$ and a nonzero sequence of coefficients $c_q$ such that for any principal character $\chi_0$ modulo $q$

$$\sum_{n<x} \alpha(n)\chi_0(n) = (c_q + o(1))x^s l(x), \quad x \to +\infty,$$

and for any nonprincipal character $\chi$

$$\sum_{n<x} \alpha(n)\chi(n) = o(x^s|l(x)|), \quad x \to +\infty.$$

In the simpler case of positive multiplicative coefficients, results of Wirsing may be applied to obtain the desired asymptotic behavior in some cases (see [Lucht 1981, Corollary 3]) but even then, the simple

*MSC2010:* primary 11N37; secondary 30B30.

*Keywords:* power series, positive completely multiplicative coefficients, completely multiplicative functions, Omega estimates.

multiplicative function defined by $\alpha_0(2^a m) = 3^a$, where $m$ is an odd number and $a$ is a nonnegative integer, cannot satisfy the conditions of Lucht's class: since

$$\sum_{n\le 2^m-1} \alpha_0(n) = 3^m - 2^m \quad \text{and} \quad \sum_{n\le 2^m} \alpha_0(n) = 2(3^m) - 2^m,$$

the function $x^{-\log 3/\log 2}\sum_{n\le x}\alpha_0(n)$ is not slowly oscillating.

The scope of this article is restricted to nonnegative, completely multiplicative functions.

An arithmetical function $\alpha(n)$ is called completely multiplicative if

$$\alpha(mn) = \alpha(m)\alpha(n)$$

for each $m$ and $n$.

For example $n^z$ is a completely multiplicative function. A Dirichlet character is also a completely multiplicative function.

Denote $e^{2\pi i\beta}$ by $e(\beta)$. Denote by $\mathfrak{A}(z)$, where $z \in \mathbb{C}$, the power series

$$\sum_{n=1}^{\infty} \alpha(n) z^n.$$

Denote by $A(x, \beta)$, where $x \in \mathbb{R}^+,\ \beta \in \mathbb{R}$, the sum

$$\sum_{n<x} \alpha(n) e(n\beta).$$

Throughout the paper the letter $p$ always denotes a generic prime number and $\sigma$ the real part of the complex number $s$. Let $g(x) > 0$. The equality $f(x) = \Omega(g(x))$ when $x \to a$ means that there is an infinite sequence $t_k \to a$ such that $|f(t_k)| > \delta g(t_k)$ for some $\delta > 0$. The relation $f(x) = \Omega(g(x))$ when $x \to a$ is also equivalent to $\overline{\lim}_{x\to a} |f(x)/g(x)| > 0$.

In [Petrushov 2018] we proved $\Omega$-estimates of power series with positive completely multiplicative coefficients.

To be specific, for a completely multiplicative function $\alpha(n)$ such that $\alpha(p) \le p$ and $0 < A \le \alpha(p) \le B < 2A$ for any prime $p$, we proved that there is a computable constant $C > 0$ such that for any $l \in \mathbb{Z}$ and any prime $q$ with $\alpha(q) \ne 1$, we have

$$\mathfrak{A}\left(e\left(\frac{l}{q}\right)r\right) = \Omega\left(\frac{|\ln(1-r)|^{C-1}}{1-r}\right)$$

as $r \to 1-$.

This implies that if $\alpha(p) \ne 1$ for infinitely many primes $p$, the natural boundary of $\mathfrak{A}(z)$ is the unit circle.

In the present article we substantially enlarge the set of completely multiplicative functions for which the associated power series has the unit circle as natural boundary.

**Theorem 1.** *Let $\alpha(n)$ be a completely multiplicative function satisfying*:

(1) $\sum_p \alpha(p)(1 - \Re(\chi(p)))/p$ *diverges for any nonprincipal* $\chi$.

(2) *The series*

$$\sum_p \frac{\alpha(p)}{p^\sigma}$$

*converges for* $\Re s > 1$.

*If the series* $\sum_{n=1}^{\infty} \alpha(n) z^n$ *has a nonsingular point on the unit circle, then* $\alpha(n) \equiv 1$.

Any completely multiplicative function $\alpha(n)$ satisfying $\alpha(p) \geq A > 0$ for every prime $p$ satisfies condition (1) of Theorem 1, and any completely multiplicative function satisfying $|\alpha(p)| \leq B$ for every prime $p$ satisfies condition (2). Therefore this theorem improves results obtained in [Petrushov 2018] and covers functions which are not in Lucht's class, such as our example $\alpha_0(n)$.

Theorem 1 is easily derived from $\Omega$-estimates for the power series $\mathfrak{A}(z)$ along every radius $[0, e(l/q))$ where every prime factor $p$ of $q$ satisfies $\alpha(p) \neq 1$.

**Theorem 2.** *Let* $\alpha(n)$ *be a positive completely multiplicative function satisfying conditions* (1) *and* (2) *of Theorem 1. Let* $q$ *be a positive integer whose prime factors all satisfy* $\alpha(p) \neq 1$. *Let* $\beta = l/q$ *with* $(l, q) = 1$.

*Set* $\delta = \sup_p(\log \alpha(p)/\log p) - 1$ *if there are primes* $p$ *with* $\alpha(p) > p$, *and* $\delta = 0$ *otherwise. Choose* $m \geq 0$ *such that there are at least* $m$ *distinct primes satisfying* $\alpha(p) = p^{1+\delta}$. *If* $\delta = 0$, *we set* $\epsilon \geq 0$ *such that there is* $c \in \mathbb{R}$ *with*

$$\sum_p \frac{\alpha(p)}{p^\sigma} \geq \epsilon |\ln(\sigma - 1)| + c \quad \text{as } \sigma \to 1+.$$

*Assume first that*

(3.1) $\delta = 0$.

*Then for all* $b < 0$

$$\mathfrak{A}(e(\beta)r) = \Omega\left(\frac{1}{(1-r)|\ln(1-r)|^{1-b}}\right) \quad \text{as } r \to 1-,$$

$$A(x, \beta) = \Omega\left(\frac{x}{(\ln x)^{1-b}}\right) \quad \text{as } x \to +\infty.$$

*If moreover*

(3.2) $\delta = 0$ *and* $\epsilon + m > 0$,

*then one can replace* $b$ *by* $\epsilon + m$ *in the previous formulas*; *that is*,

$$\mathfrak{A}(e(\beta)r) = \Omega\left(\frac{|\ln(1-r)|^{\epsilon+m-1}}{1-r}\right) \quad \text{as } r \to 1-,$$

$$A(x, \beta) = \Omega\left(x(\ln x)^{\epsilon+m-1}\right) \quad \text{as } x \to +\infty.$$

*Now*, *if*

(3.3) $\delta > 0$ *and* $\alpha(q) = q^{1+\delta}$,

*then*

$$\mathfrak{A}(e(\beta)r) = \Omega\left(\frac{|\ln(1-r)|^{m-1}}{(1-r)^{1+\delta}}\right) \quad \text{as } r \to 1-,$$

$$A(x, \beta) = \Omega\left(x^{1+\delta}(\ln x)^{m-1}\right) \quad \text{as } x \to +\infty.$$

To get these estimates, we follow a method that we developed in [Petrushov 2015b; 2017] to study power series with additive coefficients and to prove similar $\Omega$-estimates.

We use the Mellin transforms of $\mathfrak{A}(e(\beta)r)$ and $A(x,\beta)$, which both turn out be easily expressed in terms of the twisted Dirichlet series

$$F[\beta](s)=\sum_n e(\beta n)\frac{\alpha(n)}{n^s}.$$

When $\beta=l/q$ with $(l,q)=1$, $F[\beta](s)$ can be decomposed into a linear combination of

$$F(s)=\sum_n \frac{\alpha(n)}{n^s} \quad \text{and} \quad F(s,\chi)=\sum_n \frac{\alpha(n)\chi(n)}{n^s},$$

where $\chi$ runs among nonprincipal characters mod $q$. This decomposition gives a nice description of the meromorphic extension of $F[\beta](s)$ on the half-plane $\Re s>1$. Finally Tauberian arguments allow us to deduce the different $\Omega$-estimates of Theorem 2 from the corresponding analytic properties of $F[\beta](s)$.

In Section 2 we prove that the decomposition of $F[\beta](s)$ into a linear combination of $L(s,\chi)$ is possible, study some sums with characters, and study the Mellin transform. In Section 3 we prove some general $\Omega$-estimates. In Section 4 we prove the theorems and prove the generalization of Theorem 2.

## 2. Preliminary results

In this section we decompose $F[\beta](s)$ into a linear combination of $F(s,\chi)$ and prove a useful integral equality.

Let $q$ be a natural number, $q>1$ and let $q=\prod_{i=1}^k p_i^{l_i}$ be its decomposition into prime factors throughout this section. Let $K(q)=\left\{n\in\mathbb{N}: n=\prod_{i=1}^k p_i^{m_i}\right\}$ and in this definition $m_i$ are arbitrary nonnegative integers. From the fundamental theorem of arithmetic it easily follows that each $n\in\mathbb{N}$ has a unique representation

$$n=km, \tag{1}$$

where $k\in K(q)$, $(m,q)=1$.

For a Dirichlet character $\chi$ modulo $q$, we denote by $\tau(\chi,l)$ the Gauss sum

$$\sum_{n=1}^{q}\chi(n)e\left(\frac{nl}{q}\right).$$

We define $C_\chi(s)$ as the Dirichlet series

$$\sum_{k\in K(q)}\frac{\alpha(k)}{k^s}\tau(\bar{\chi},lk),$$

where $\chi$ is the conjugate character of $\chi$. Throughout the paper $C_\chi(s)$ depends on $l$ and $q$.

**Lemma 3.** *Let $\alpha(n)$ be a multiplicative function and let the Dirichlet series*

$$F(s)=\sum_{n=1}^{\infty}\frac{\alpha(n)}{n^s}$$

*be absolutely convergent in the domain* $\{\Re s > \sigma_1\}$. *Then for* $\Re s > \sigma_1$ *the following identity holds*:

$$F[\beta](s) = \frac{1}{\phi(q)} \sum_{\chi \pmod q} C_\chi(s) F(s, \chi) \tag{2}$$

(see [Petrushov 2015a, p. 20]).

Let $c_q(n)$ be Ramanujan sum,

$$c_q(n) = \sum_{\substack{0 \le l < q \\ (l,q)=1}} e\left(\frac{nl}{q}\right).$$

Recall standard properties of the Ramanujan sum:

(1) If $(q_1, q_2) = 1$ then $c_{q_1 q_2}(n) = c_{q_1}(n) c_{q_2}(n)$.

(2) $c_q(k)$ depends only on $(k, q)$.

(3) $c_k(q) = \tau(\chi_0, k)$, where $\chi_0$ is the principal character modulo $q$.

**Lemma 4.** *Let* $f$ *be a completely multiplicative function with* $|f(p)| < 1$ *for* $p \mid q$. *We have*

$$\left(\sum_{n \in K(q)} f(n) c_q(n)\right) \prod_{p \mid q} (1 - f(p)) = \prod_{p^m \,||\, q} ((pf(p))^m - (pf(p))^{m-1}), \tag{3}$$

*where the notation* $p^m \,||\, q$ *means that the multiplicity of* $p$ *in the prime decomposition of* $q$ *is* $m$.

*Proof.* We see

$$\begin{aligned}
\left(\sum_{n \in K(q)} f(n) c_q(n)\right) \prod_{p \mid q} (1 - f(p)) &= \prod_{p \mid q} \left(\sum_{n \in K(p)} f(n) c_q(n)\right) (1 - f(p)) \\
&= \prod_{p^m \,||\, q} \left(\sum_{j=0}^{\infty} f^j(p) c_{p^j}(p^m)\right) (1 - f(p)) \\
&= \prod_{p^m \,||\, q} \left(-p^{m-1} f^{m-1}(p) + p^{m-1}(p-1) \frac{f^m(p)}{1 - f(p)}\right) (1 - f(p)) \\
&= \prod_{p^m \,||\, q} (-p^{m-1} f^{m-1}(p)(1 - f(p) + p^{m-1}(p-1) f^m(p)) \\
&= \prod_{p^m \,||\, q} (-p^{m-1} f^{m-1}(p) + p^{m-1} f^m(p) + p^m f^m(p) - p^{m-1} f^m(p)) \\
&= \prod_{p^m \,||\, q} ((pf(p))^m - (pf(p))^{m-1}).
\end{aligned}$$

□

**Lemma 5.** *Let* $\alpha(n)$ *be a completely multiplicative function, and let* $\sigma_1$ *be as in Lemma 3. The following formula holds*:

$$C_{\chi_0}(s) F(s, \chi_0) = \prod_{i=1}^{r} \left(\frac{p_i^{m_i - 1} \alpha^{m_i - 1}(p_i)}{p_i^{(m_i - 1)s}} (\alpha(p_i) p_i^{1-s} - 1)\right) F(s). \tag{4}$$

*Proof.* If the series $\sum_n |\alpha(n)|/n^\sigma$ converges for $\sigma > \sigma_1$, then $|\alpha(p)| \le p^{\sigma_1}$ for all prime $p$. The completely multiplicative function defined by $f(p) = \alpha(p)/p^s$ satisfies the condition of Lemma 4 if $\Re s > \sigma_1$. Along with the facts that $\tau(\chi_0, l) = c_q(l)$ and $F(s, \chi_0) = \prod_{p \mid q}(1 - f(p)/p^s)F(s)$, this lemma follows from Lemma 4 with $f(p) = \alpha(p)/p^s$. □

**Lemma 6.** *Each character modulo $q$ can be expressed in the form*

$$\chi = \chi_1 \chi_2,$$

*where $\chi_1$ is the principal character modulo $q_1$, $\chi_2$ is a character induced by a primitive character modulo $q_2^{'}$, $q_2^{'} \mid q_2$, $q = q_1 q_2$, $(q_1, q_2) = 1$, and the prime divisors of $q_2$ and $q_2^{'}$ are the same.*

Lemma 6 follows from [Apostol 1976, Theorem 8.18, p. 171].

**Lemma 7.** *Let $q = q_1 q_2$, $(q_1, q_2) = 1$. Let $\chi = \chi_1 \chi_2$, where $\chi_1$ and $\chi_2$ are characters modulo $q_1$ and $q_2$ respectively. Then for each $l \in \mathbb{Z}$ we have $\tau(\chi, l) = \chi_2(q_1)\chi_1(q_2)\tau(\chi_1, l)\tau(\chi_2, l)$.*

Lemma 7 follows from [Montgomery and Vaughan 2007, Theorem 9.6, p. 287].

**Lemma 8.** *Let $\chi$ be a character modulo $q = \prod_i p_i^{l_i}$ induced by a primitive character modulo $q_2^{'}$, and suppose $q_2^{'} \mid q_2$, $q = q_1 q_2$, $(q_1, q_2) = 1$, and the prime divisors of $q_2$ and $q_2^{'}$ are the same. If there is an $i$ such that $p_i^{l_i}$ divides $m$ then $\tau(\chi, m) = 0$.*

*Proof.* Let $m = \prod_{i=1}^r p_i^{n_i}$. Let $\chi = \prod_{i=1}^r \chi_i$, where $\chi_i$ are characters induced by characters modulo $p_i^{r_i}$. Then by Lemma 7 and that fact that $\tau(\chi, al) = \bar{\chi}(a)\tau(\chi, l)$ (see [Montgomery and Vaughan 2007, Theorem 9.5, p. 287]), we obtain

$$\tau(\chi, lk) = C \prod_{i=1}^{r} \tau(\chi_i, p_i^{n_i}),$$

where $|C| = 1$. Since $p_i^{l_i} \mid m$ we have $n_i \ge l_i \ge r_i$. Thus $\tau(\chi_i, p_i^{n_i}) = 0$. □

**Lemma 9.** *Let $\Re s > \sigma_1$, where $\sigma_1$ is as defined in Lemma 3. Let*

$$\chi = \chi_1 \chi_2,$$

*where $\chi_1$ is the principal character modulo $q_1$, $\chi_2$ is a character induced by a primitive character modulo $q_2^{'}$, $q_2^{'} \mid q_2$, $q = q_1 q_2$, $(q_1, q_2) = 1$, and the prime divisors of $q_2$ and $q_2^{'}$ are the same. Then*

$$C_\chi(s) = \bar{\chi}_2(q_1)\bar{\chi}_2(l) \sum_{k_1 \in K(q_1)} \frac{\tau(\chi_1, k_1)}{k_1^s} \chi_2(k_1)\alpha(k_1) \sum_{k_2 \mid q_2} \frac{\tau(\chi_2, k_2)}{k_2^s} \alpha(k_2). \tag{5}$$

*Proof.* Using Lemmas 6 and 7 and the fact that $\tau(\chi, al) = \bar{\chi}(a)\tau(\chi, l)$ when $\chi(a) \ne 0$ we derive

$$C_\chi(s) = \sum_{\substack{k_1 \in K(q_1) \\ k_2 \in K(q_2)}} \frac{\tau(\chi_1 \bar{\chi}_2, lk_1k_2)}{k_1^s k_2^s} \alpha(k_1)\alpha(k_2) = \sum_{\substack{k_1 \in K(q_1) \\ k_2 \in K(q_2)}} \frac{\bar{\chi}_2(q_1)\tau(\chi_1, lk_1k_2)\tau(\bar{\chi}_2, lk_1k_2)\alpha(k_1)\alpha(k_2)}{k_1^s k_2^s}$$

$$= \bar{\chi}_2(q_1)\bar{\chi}_2(l) \sum_{k_1 \in K(q_1)} \frac{\tau(\chi_1, k_1)}{k_1^s} \chi_2(k_1)\alpha(k_1) \sum_{k_2 \in K(q_2)} \frac{\tau(\chi_2, k_2)}{k_2^s} \alpha(k_2). \tag{6}$$

By Lemma 8 for $k_2 \in K(q_2)$ we have $\tau(\chi_2, k_2) = 0$ unless $k_2 \mid q_2$. Hence the second sum may be written as a Dirichlet polynomial

$$\sum_{k_2 \mid q_2} \frac{\tau(\chi_2, k_2)}{k_2^s} \alpha(k_2). \qquad \square$$

**Lemma 10.** *Let*

$$\chi = \chi_1 \chi_2,$$

*where $\chi_1$ is the principal character modulo $q_1$, $\chi_2$ is character induced by a primitive character modulo $q_2'$, $q_2' \mid q_2$, $q = q_1 q_2$, $(q_1, q_2) = 1$, and the prime divisors of $q_2$ and $q_2'$ are the same. Then*

$$C_\chi(s) F(s, \chi) = A_\chi(s) F(s, \chi_2),$$

*where $A_\chi(s)$ is an entire function.*

*Proof.* Let $q = \prod_{i=1}^r p_i^{l_i}$. From Lemma 9 we get

$$C_\chi(s) F(s, \chi) = B_\chi(s) \sum_{k_1 \in K(q_1)} \frac{\tau(\chi_1, k_1)}{k_1^s} \chi_2(k_1) \alpha(k_1) \prod_{i=1}^r \left(1 - \frac{\chi_2(p_i)\alpha(p_i)}{p_i^s}\right) F(s, \chi_2),$$

where $B_\chi(s)$ is a Dirichlet polynomial. Hence

$$C_\chi(s) F(s, \chi) = B_\chi(s) \sum_{k_1 \in K(q_1)} \frac{\alpha(k_1)}{k_1^s} \chi_2(k_1) c_{q_1}(k_1) \prod_{i=1}^r \left(1 - \frac{\chi_2(p_i)}{p_i^s} \alpha(p_i)\right) F(s, \chi_2). \tag{7}$$

Using Lemma 4 with $f(p) = \alpha(p)\chi_2(p)/p^s$, we obtain

$$C_\chi(s) F(s, \chi) = B_\chi(s) \prod_{i=1}^r \left(\left(p \frac{\chi_2(p)\alpha(p)}{p^s}\right)^{l_i} - \left(p \frac{\chi_2(p)\alpha(p)}{p^s}\right)^{l_i - 1}\right) F(s, \chi_2) = A_\chi(s) F(s, \chi_2),$$

where $A_\chi(s)$ is a Dirichlet polynomial. $\square$

**Lemma 11.** *The following equality holds*:

$$F[\beta](s) = \frac{1}{\phi(q)} \sum_{\chi \in X} D_\chi(s) F(s, \chi) + \frac{1}{\phi(q)} D(s) F(s), \tag{8}$$

*where $X$ is the set of primitive characters of modulus $q_1$ with $q_1 \mid q$,*

$$D(s) = \prod_{i=1}^r \left( \frac{p_i^{l_i - 1} \alpha^{l_i - 1}(p_i)}{p_i^{(l_i - 1)s}} (\alpha(p_i) p_i^{1-s} - 1)\right), \tag{9}$$

*and $D_\chi(s)$ are Dirichlet polynomials just like $D(s)$.*

*Proof.* The proof follows from Lemmas 3, 5, 6 and 10. $\square$

Note that $D(1) = 0$ if and only if $\alpha(p) = 1$ for some $p \mid q$.

The following proposition relates $F[l/q](s)$ to $\mathfrak{A}(e(l/q)r)$ and $A(x, \beta)$.

**Proposition 12.** *Let $\alpha(n)$ be a sequence such that $|\alpha(n)| \le n^{1+\delta}$. Then for any $\beta \in \mathbb{R}$, for $\Re s > 2+\delta$*

$$\Gamma(s)F[\beta](s) = \int_0^\infty t^{s-1}\mathfrak{A}(e(\beta)e^{-t})\,dt,$$

$$\frac{1}{s}F[\beta](s) = \int_0^1 t^{s-1}A(t^{-1},\beta)\,dt.$$

*Proof.* From inequality $|\alpha(n)| \le n^{1+\delta}$ we derive $A(x,0) \ll x^{2+\delta}$ and

$$|\mathfrak{A}(e^{-t})| \le \sum_{n=1}^{+\infty} n^{1+\delta}e^{-nt} \le e^t \int_0^{+\infty} u^{1+\delta}e^{-ut}\,dt \le e^t\Gamma(2+\delta)t^{-2-\delta}.$$

Using the Lebesgue dominated convergence theorem with $|\mathfrak{A}(e(\beta)e^{-t})| \le |\mathfrak{A}(e^{-t})|$ we obtain

$$\int_0^\infty t^{s-1}\mathfrak{A}(e(\beta)e^{-t})\,dt = \sum_{n=1}^\infty \alpha(n)e(\beta n)\int_0^\infty t^{s-1}e^{-nt}\,dt = \sum_{n=1}^\infty \alpha(n)e(\beta n)\Gamma(s)n^{-s}.$$

Let $B_m(x) = 1$ if $x < 1/m$ and $B_m(x) = 0$ if $x \ge 1/m$. Then $\sum_m \alpha(m)B_m(x) = A(1/x,0) \ll x^{-2-\delta}$. Using the Lebesgue dominated convergence theorem with $|A(1/t,\beta)| \le |A(1/t,0)|$ we obtain

$$\int_0^1 x^{s-1}\sum_{m=1}^\infty \alpha(m)e(m\beta)B_m(x)\,dx = \sum_{m=1}^\infty \alpha(m)e(m\beta)\int_0^1 x^{s-1}B_m(x)\,dx = \sum_{m=1}^\infty \alpha(m)e(m\beta)\int_0^{1/m} x^{s-1}\,dx$$

$$= \frac{1}{s}\sum_{m=1}^\infty \alpha(m)e(m\beta)m^{-s} = \frac{1}{s}F[\beta](s). \qquad \square$$

## 3. Growth of some functions

In this section we study growth of some Euler products as $s \to 1$.

Let

$$G(s) = \prod_{\alpha(p)\le p}\left(1-\frac{\alpha(p)}{p^s}\right)^{-1} \quad \text{and} \quad H(s) = \prod_{\alpha(p)>p}\left(1-\frac{\alpha(p)}{p^s}\right)^{-1}.$$

Throughout this section, we assume that condition (2) of Theorem 1 is satisfied, that is, we assume the convergence of the series $\sum_p \alpha(p)/p^\sigma$ for any $\sigma > 1$. It follows that $G(s)$ is an analytic function on $\{\Re s > 1\}$. It also follows that for any $\epsilon > 0$ there are at most finitely many $p$ such that $\alpha(p) > p^{1+\epsilon}$.

**Lemma 13.** *Let $\chi$ be a nonprincipal character modulo $q$. If the series*

$$\sum_p \alpha(p)\frac{1-\Re\chi(p)}{p^{1+x}}$$

*diverges, then the following relation holds*:

$$G(1+x,\chi) = o(G(1+x)) \quad \text{as } x \to 0+. \tag{10}$$

*Proof.* The condition (10) is equivalent to

$$\ln|G(1+x)| - \ln|G(1+x,\chi)| \to +\infty$$

as $x \to 0+$. Let $x > 0$. We have

$$\ln|G(1+x)| - \ln|G(1+x,\chi)| = -\sum_{\alpha(p)\le p} \ln(1-\alpha(p)p^{-1-x}) + \sum_{\alpha(p)\le p} \ln|1-\alpha(p)\chi(p)p^{-1-x}|$$
$$= \sum_{\alpha(p)\le p} -\ln(1-\alpha(p)p^{-1-x}) + \sum_{\alpha(p)\le p} \Re\ln(1-\alpha(p)\chi(p)p^{-1-x}),$$

where ln is the principal value of the logarithm.

Using the power series expression of $-\ln(1-z)$ on the unit disk we have

$$\ln|G(1+x)| = -\sum_{\alpha(p)\le p} \ln(1-\alpha(p)p^{-1-x}) = \sum_{\alpha(p)\le p}\sum_{k\ge 1} \frac{1}{k}\alpha(p)^k p^{-k(1+x)}.$$

Notice that the summands are nonnegative and the double sum converges. Similarly

$$\ln|G(1+x,\chi)| = -\sum_{\alpha(p)\le p} \Re\ln(1-\alpha(p)p^{-1-x}\chi(p)) = \sum_{\alpha(p)\le p}\sum_{k\ge 1} \Re\frac{1}{k}(\chi^k(p)\alpha(p)^k p^{-k(1+x)}),$$

where the double sum is absolutely convergent (using the previous double sum). Therefore

$$\ln|G(1+x)| - \ln|G(1+x,\chi)| = \sum_{\alpha(p)\le p}\sum_{k>1} \frac{1}{k}(1-\Re\chi^k(p))\alpha(p)^k p^{-k(1+x)} \ge \sum_{\alpha(p)\le p} (1-\Re\chi(p))\alpha(p)p^{-(1+x)}$$

since all summands are nonnegative. If $\sum_p \alpha(p)(1-\Re\chi(p))/p^{1+x}$ is bounded as $x$ tends to 0 then it has a limit and by the Tauberian theorem (see [Hardy and Littlewood 1914, Theorem 17]) the sum $\sum_p \alpha(p)(1-\Re\chi(p))/p$ is convergent, which contradicts our assumption. Therefore we have

$$\Re\sum_p \frac{\alpha(p)}{p^{1+x}}(1-\chi(p)) \to +\infty$$

for $x \to 0+$. □

**Lemma 14.** *Let assumption* (2) *of Theorem 1 hold. Then*

$$|F(1+x,\chi)| \le \prod_{p\mid q} |1-\alpha(p)p^{-1-x}|\,|F(1+x)|,$$

*where $q$ is the modulus of $\chi$.*

*Proof.* For any $t \in [0,+\infty)$ and $z \in \mathbb{C}$ with $|z| = 1$, we have $|t-1| \le |t-z|$(since 1 is the point of the unit circle closest to any given point of $[0,\infty)$) or again $|t-1| \le |tz-1|$. For $p$ with $\alpha(p) \ge 0$ and $\chi(p) \ne 0$ we have $|1-\alpha(p)\chi(p)p^{-1-x}|^{-1} \le |1-\alpha(p)p^{-1-x}|^{-1}$. Hence we deduce

$$|F(1+x,\chi)| = \prod_{p\nmid q} |1-\alpha(p)\chi(p)p^{-1-x}|^{-1}$$
$$\le \prod_{p\nmid q} |1-\alpha(p)p^{-1-x}|^{-1} = \prod_{p\mid q} |1-\alpha(p)p^{-1-x}|\,|F(1+x)|.$$ □

Let $f$ be locally integrable on $(0, +\infty)$. The Mellin transform of $f$ is defined by the integral $f^*(s) = \int_0^\infty x^{s-1} f(x)\,dx$. The fundamental strip is the largest open strip on which it is defined. In particular, if $f$ satisfies the asymptotic conditions $f(x) = O(x^{-\sigma})$ as $x \to 0+$ for some $\sigma \in \mathbb{R}$ and $f(x) = o(x^{-N})$ as $x \to +\infty$ for any $N > 0$, then the integral defining $f^*(s)$ is convergent for any $s \in \mathbb{C}$ such that $\Re s > \sigma$, and the Mellin transform $f^*$ is an analytic function over $\{\Re s > \sigma\}$.

**Proposition 15.** *Let $f$ be a locally integrable function on $(0, +\infty)$ such that its Mellin transform $f^*$ is analytic on $\{\Re s > \sigma_0\}$ with $\sigma_0 > 0$. Let $t_0 \in \mathbb{R}$. If $\overline{\lim}_{\sigma \to \sigma_0+} |f^*(\sigma + it_0)| = +\infty$, then for any $b > 0$, we have*

$$\overline{\lim_{x \to 0+}} \frac{|f(x)|}{x^{-\sigma_0} |\ln x|^{b-1}} = +\infty.$$

*Moreover, if there are $b > 0$ and $c > 0$ such that*

$$\overline{\lim_{\sigma \to \sigma_0+}} \frac{|f^*(\sigma + it_0)|}{(\sigma - \sigma_0)^b} \geq c,$$

*then*

$$\overline{\lim_{x \to 0+}} \frac{|f(x)|}{x^{-\sigma_0} |\ln x|^{b-1}} \geq \frac{c}{\Gamma(b)}.$$

*Proof.* Since $f$ admits a Mellin transform, there exists $\sigma_1 > \sigma_0$ such that $\int_0^{+\infty} u^{\sigma_1 - 1} |f(u)|\,du$ is convergent. Therefore, for any $s$ such that $\Re s < \sigma_1$, we have

$$\left| \int_1^{+\infty} u^{s-1} f(u)\,du \right| \leq \int_1^{+\infty} u^{\sigma_1 - 1} |f(u)|\,du \leq |f|^*(\sigma_1).$$

Assume now there are constants $b \in \mathbb{R}$, $c > 0$, and $u_0 \in (0, 1)$ such that $|f(u)| \leq c' u^{-\sigma_0} |\ln u|^{b-1}$ for any $u \leq u_0$. In that case, for any $s$ satisfying $\Re s > \sigma_0$, the function $u^{s-1} f(u)$ is integrable on $(0, 1)$ and

$$\left| \int_0^1 u^{s-1} f(u)\,du \right| \leq c' \int_0^{u_0} u^{\sigma - \sigma_0 - 1} |\ln u|^{b-1}\,du + \int_{u_0}^1 u^{\sigma_0 - 1} |f(u)|\,du.$$

We deduce that for any $s$ such that $\sigma_0 < \Re s \leq \sigma_1$, we have

$$|f^*(s)| \leq c' \int_0^{u_0} u^{\sigma - \sigma_0} |\ln u|^{b-1}\,du + C,$$

where $C$ is some constant independent of $s$.

If we can choose $b < 0$, then the integral $\int_0^{u_0} u^{\sigma - \sigma_0 - 1} |\ln u|^{b-1}\,du$ is bounded by the convergent integral $\int_0^{u_0} u^{-1} |\ln u|^{b-1}\,du$ which does not depend on $s$. We conclude that $\overline{\lim}_{\sigma \to \sigma_0+} |f^*(\sigma + it_0)| < +\infty$ for any $t_0$. This proves the first case.

If $b > 0$, then

$$\int_0^1 u^{(\sigma - \sigma_0) - 1} |\ln u|^{b-1}\,du = \int_0^{+\infty} e^{-(\sigma - \sigma_0)u} u^{b-1}\,du = \Gamma(b)(\sigma - \sigma_0)^b.$$

We conclude that $|f^*(s)| \leq c' \Gamma(b)(\sigma - \sigma_0)^b + C$ and consequently

$$\overline{\lim_{\sigma \to \sigma_0+}} \frac{|f^*(\sigma + it_0)|}{(\sigma - \sigma_0)^{-b}} \leq c' \Gamma(b).$$

Hence, if this limit superior is larger than $c$, it contradicts the previous conclusion for any $c' < c/\Gamma(b)$ and the assumption $|f(u)| \le c' u^{\sigma_0} |\ln u|^{b-1}$ has to be contradicted in any neighborhood of 0 and for any $c' < c/\Gamma(b)$. This implies

$$\varlimsup_{x\to 0+} \frac{|f(x)|}{x^{-\sigma_0}|\ln x|^{b-1}} \ge \frac{c}{\Gamma(b)}. \qquad \square$$

**Proposition 16.** *Let $\alpha(n)$ be a sequence satisfying $|\alpha(n)| \le n^{1+\delta}$. Let $\beta \in \mathbb{R}$ and $\sigma_0 > 0$ such that the Dirichlet series $F[\beta](s) = \sum_n e(n\beta)\alpha(n)n^{-s}$ is analytic on $\{\Re s > \sigma_0\}$. If*

$$\varlimsup_{\sigma\to\sigma_0+} |F[\beta](\sigma + it_0)| = +\infty,$$

*then for any $b > 0$*

$$\mathfrak{A}(e(\beta)r) = \Omega((1-r)^{-\sigma_0}|\ln(1-r)|^{b-1}),$$
$$A(x,\beta) = \Omega(x^{\sigma_0}(\ln x)^{b-1}).$$

*Let*

$$\varlimsup_{\sigma\to\sigma_0+} \frac{|F[l/q](\sigma + it_0)|}{(\sigma-\sigma_0)^{-b}} \ge C.$$

*Then*

$$\mathfrak{A}(e(\beta)r) = \Omega((1-r)^{-\sigma_0}|\ln(1-r)|^{b-1}),$$
$$A(x,\beta) = \Omega(x^{\sigma_0}(\ln x)^{b-1}).$$

*Proof.* The assumptions of Proposition 12 are satisfied. Therefore $\Gamma(s)F[\beta](s)$ and $(1/s)F[\beta](s)$ are the Mellin transforms of $\mathfrak{A}(e(\beta)e^{-t})$ and $A(t^{-1},\beta)\mathbf{1}_{(0,1)}(t)$ and they satisfy the assumptions of Proposition 15.

By change of variable $r = e^{-u}$ we have

$$\varlimsup_{u\to 0+} \frac{\mathfrak{A}(e(\beta)e^{-u})}{u^{-\sigma_0}|\ln u|^{b-1}} = \varlimsup_{r\to 1-} \frac{\mathfrak{A}(e(\beta)r)}{(1-r)^{-\sigma_0}|\ln(1-r)|^{b-1}},$$

and by change of variable $y = x^{-1}$ we have

$$\varlimsup_{x\to 0+} \frac{|A(x^{-1},\beta)|}{x^{-\sigma_0}|\ln x|^{b-1}} = \varlimsup_{y\to+\infty} \frac{|A(y,\beta)|}{y^{\sigma_0}(\ln y)^{b-1}}.$$

Consider the first case. We see by Proposition 12

$$f^*(s) = \Gamma(s)F\left[\frac{l}{q}\right](s) = \int_0^\infty t^{s-1}\mathfrak{A}\left(e\left(\frac{l}{q}\right)e^{-t}\right)dt,$$
$$\varlimsup_{\sigma\to\sigma_0+} |f^*(\sigma + it_0)| = +\infty.$$

Thus by Proposition 15

$$\varlimsup_{u\to 0+} \frac{\mathfrak{A}(e(l/q)e^{-u})}{u^{-\delta}|\ln u|^{b-1}} = +\infty$$

for each $b > 0$.

Similarly

$$f^*(s) = \int_0^\infty x^{-s-1} A(x^{-1}, \beta)\, dx = \frac{1}{s} F\left[\frac{l}{q}\right](s),$$
$$\overline{\lim_{\sigma \to \sigma_0+}} |f^*(\sigma + it_0)| = +\infty.$$

Thus by Proposition 15

$$\overline{\lim_{x \to 0+}} \frac{|A(x^{-1}, \beta)|}{x^{-\sigma_0} |\ln x|^{b-1}} = +\infty \quad \text{and} \quad \overline{\lim_{y \to +\infty}} \frac{|A(y, \beta)|}{y^{\sigma_0} (\ln y)^{b-1}} = +\infty$$

for each $b > 0$.

Consider the second case. Let

$$\overline{\lim_{\sigma \to \sigma_0+}} \frac{|F[l/q](\sigma + t_0)|}{(\sigma - \sigma_0)^{-b}} \geq c.$$

Thus by Proposition 15

$$\overline{\lim_{u \to 0+}} \frac{|\mathfrak{A}(e(l/q)e^{-u})|}{u^{-\delta} |\ln u|^{-1+b}} \geq \frac{c}{\Gamma(b)}.$$

Further

$$\overline{\lim_{\sigma \to \sigma_0+}} \frac{(\sigma + it_0) F(\sigma + it_0)}{(\sigma - \sigma_0)^{-b}} \geq c \frac{1}{|\sigma_0 + it_0|}.$$

Hence by Proposition 15

$$\overline{\lim_{x \to 0+}} \frac{A(x^{-1}, \beta)}{x^{-\sigma_0} |\ln x|^{-1+b}} \geq \frac{c}{|\sigma_0 + it_0| \Gamma(b)} \quad \text{and} \quad \overline{\lim_{y \to +\infty}} \frac{|A(y, \beta|)}{y^{\sigma_0} (\ln y)^{-1+b}} \geq \frac{c}{|\sigma_0 + it_0| \Gamma(b)}. \qquad \square$$

By Lemma 11

$$F[\beta](s) = \frac{1}{\phi(q)} \sum_{\chi \in X} D_\chi(s) F(s, \chi) + \frac{1}{\phi(q)} D(s) F(s),$$

where $X$ is the set of primitive characters of modulus $q_1$ with $q_1 \mid q$, $D_\chi(s)$ are entire functions, $D(s)$ is defined in (9).

**Proposition 17.** *Let $\alpha(n)$ be a positive completely multiplicative function satisfying conditions* (1) *and* (2) *of Theorem 1. Let $\beta = l/q$ with $(l, q) = 1$.*

*If $\alpha(p) \leq p$ for all primes $p$, and $\alpha(p) \neq 1$ for all primes $p$ dividing $q$, then*

$$F[\beta](1 + x) \sim \frac{D(1)}{\phi(q)} F(1 + x) \quad \text{as } x \to 0+.$$

*If there exists $\delta > 0$ such that $\alpha(p) = p^{1+\delta}$ for every prime factor $p$ of $q$, then*

$$F[\beta](1 + x) \sim F(1 + x) \quad \text{as } x \to \delta.$$

*Proof.* Consider the first case. Since $D_\chi(s)$ is entire we see

$$|D_\chi(1 + x) F(1 + x, \chi)| \ll |F(1 + x, \chi)|$$

as $x \to 0+$. Further by Lemma 13 we obtain (in our case $F = G$)

$$|D_\chi(1+x)F(1+x,\chi)| \ll |F(1+x,\chi)| = o(F(1+x))$$

as $x \to 0+$. Hence for each $\chi \neq \chi_0 \pmod q$ we have $D_\chi(1+x)F(1+x,\chi) = o(F(1+x))$ as $x \to 0+$. Thus

$$\frac{1}{\phi(q)} \sum_{\chi \in X} D_\chi(1+x)F(1+x,\chi) = o(F(1+x)) \quad \text{as } x \to 0+.$$

It follows that

$$F[\beta](1+x) \sim \frac{1}{\phi(q)} D(1+x)F(1+x) \quad \text{as } x \to 0+.$$

From the expression (9) for $D(s)$, we get

$$D(1) = \prod_{p^m \,||\, q} \alpha(p^{m-1})(\alpha(p) - 1).$$

Since $\alpha(p) \neq 1$ for any $p$ dividing $q$, we have $D(1) \neq 0$ and

$$F[\beta](1+x) \sim \frac{1}{\phi(q)} D(1)F(1+x) \quad \text{as } x \to 0+. \tag{11}$$

Consider the second case. Since $D_\chi(s)$ are Dirichlet polynomials we have

$$D_\chi(s)F(1+x,\chi) \ll |F(1+x,\chi)| \quad \text{as } x \to \delta$$

for any character $\chi \in X$. By Lemma 14, we have

$$|F(1+x,\chi)| \le \prod_{p \,|\, q_0} \left|1 - \frac{\alpha(p)}{p^{1+x}}\right| |F(1+x)|,$$

where $q_0$ is the modulus of $\chi$. Since $\alpha(p) = p^{1+\delta}$ for any prime $p$ dividing $q_0$, the product tends to 0 as $x \to \delta$ and

$$F(1+x,\chi) = o(|F(1+x)|) \quad \text{as } x \to \delta.$$

Again we derive that

$$F[\beta](1+x) \sim \frac{1}{\phi(q)} D(1+x)F(1+x) \quad \text{as } x \to \delta.$$

From the expression (9) for $D(s)$, we get

$$D(1+\delta) = \prod_{p^m \,||\, q} \alpha(p^{m-1}) p^{-\delta(m-1)} (\alpha(p) p^{-\delta} - 1)$$

and since $\alpha(p) = p^{1+\delta}$ for every $p \,|\, q$, we have $D(1+\delta) = \phi(q)$. Therefore

$$F[\beta](1+x) \sim F(1+x) \quad \text{as } x \to \delta. \qquad \square$$

## 4. Proof of the theorems

*Proof of Theorem 2.* Assume first that assumptions (1) and (2) of Theorem 1 and assumption (3.1) of Theorem 2 are satisfied. Under these conditions, we have established that $F[\beta](s)$ is analytic on $\{\Re s > 1\}$ and $\sum_p \alpha(p)p^{1+x} \to +\infty$ as $x \to 0+$. We derive that

$$\ln F(1+x) = -\sum_p \ln\left(1 - \frac{\alpha(p)}{p^{1+x}}\right) \geq \sum_p \frac{\alpha(p)}{p^{1+x}} \to +\infty$$

as $x \to 0+$. By Proposition 17 we obtain $F[\beta](1+x) \to \infty$. By Proposition 16 with $\sigma_0 = 1$ we obtain the $\Omega$-estimates.

Assume now that assumptions (1) and (2) of Theorem 1 and assumption (3.2) are satisfied. Under these conditions, we have established that $F[\beta](s)$ is analytic on $\{\Re s > 1\}$ and $\sum_p \alpha(p)/p^{1+x} \to +\infty$ as $x \to 0+$. We choose a set $P_0$ of $m \geq 0$ primes $p$ satisfying $\alpha(p) = p$. The Euler product $F_0(s) = \prod_{p \notin P_0}(1 - \alpha(p)p^{-s})^{-1}$ is convergent for any $s$ with $\Re s > 1$ and we have

$$\ln F_0(1+x) = -\sum_{p \in P_0} \ln\left(1 - \frac{\alpha(p)}{p^{1+x}}\right) \geq \sum_{p \in P_0} \frac{\alpha(p)}{p^{1+x}} \geq -\epsilon \ln x + c - m$$

as $x \to 0+$. On the other hand, for any $p \in P_0$ we have

$$\left(1 - \frac{\alpha(p)}{p^{1+x}}\right)^{-1} = (1 - e^{-x \ln p})^{-1} \geq \frac{1}{x \ln p}.$$

We derive that

$$F(1+x) \geq \frac{e^{c-m}}{\prod_{p \in P_0} \ln p} x^{-(\epsilon+m)}$$

when $x$ is sufficiently close to 0. By Proposition 17 we obtain

$$F[\beta](1+x) \geq C \frac{e^{c-m}}{\prod_{p \in P_0} \ln p} x^{-(\epsilon+m)} \to \infty.$$

By Proposition 16 with $\sigma_0 = 1$ we obtain the $\Omega$-estimates.

Assume now that assumptions (1) and (2) of Theorem 1 and assumption (3.3) of Theorem 2 are satisfied. Under these conditions, we have established that $F[\beta](s)$ is meromorphic over $\Re s > 1$, and that for any $\epsilon > 0$ there are finitely many primes $p$ such that $\alpha(p) = p^{1+\epsilon}$ is finite and nonempty. We can set $m$ as its cardinal. We deduce also that there exists $\delta_0 < \delta$ such that the Euler product $F_0(s) = \prod_{p \in P_0}(1 - \alpha(p)p^{-s})^{-1}$ converges for any $s$ with $\Re s > \delta_0$. In particular, $F_0(1+\delta)$ is well-defined. On the other hand, for any $p \in P_0$ we have

$$\left(1 - \frac{\alpha(p)}{p^{1+x}}\right)^{-1} = (1 - e^{(x-\delta)\ln p})^{-1} \sim \frac{1}{(x-\delta)\ln p}$$

as $x \to \delta$. We derive

$$F(1+x) \sim \frac{F_0(1+\delta)}{\prod_{p \in P_0} \ln p} (x-\delta)^{-m}$$

as $x \to \delta$. By Proposition 17 we obtain

$$F[\beta](1+x) \sim C(x-\delta)^{-m} \to \infty$$

as $x \to \delta$. By Proposition 16 with $\sigma_0 = 1+\delta$ we obtain the $\Omega$-estimates. □

*Proof of Theorem 1.* Using Theorem 2 it has been proven that the point $e(l/p^m)$ is a singular point of $\mathfrak{A}(z)$ for $l$ with $(l, p) = 1$ and $m \geq 0$, if $p$ is chosen such that $\alpha(p) \neq 1$ if $\delta = 0$ or $\alpha(p) = p^{1+\delta}$ if $\delta > 0$. This provides a dense set of singular points on the unit circle, unless $\alpha(p) = 1$ for any $p$. □

Now we can prove the generalization of Theorem 1.

**Theorem 18.** *Let $\alpha(n)$ be a positive completely multiplicative function and $\sigma_0 \in \mathbb{R}$ such that*:

(1) *For any nonprincipal Dirichlet character $\chi$ the following series diverges*:

$$\sum_p \alpha(p) \frac{1 - \Re\chi(p)}{p^{\sigma_0}}.$$

(2) *For any $\sigma > \sigma_0$, the following series converges*:

$$\sum_p \frac{\alpha(p)}{p^{\sigma}}.$$

*If the series $\sum_{n=1}^{\infty} \alpha(n) z^n$ has a nonsingular point on the unit circle, then $\alpha(n) = n^{\sigma_0 - 1}$.*

Note that the series $\sum_{n=1}^{\infty} n^s z^n$ is the polylogarithm of order $s$ which admits an analytical extension beyond the unit circle for any $s \in \mathbb{C}$.

*Proof.* We can reduce to the case where $\sigma_0 > 0$: If the series $\sum_{n=1}^{\infty} \alpha(n) z^n$ has a nonsingular point on the unit circle, that is, an analytical extension beyond the unit circle, its derivative has the same radius of convergence and the same domain of analyticity. Therefore we can apply the theorem to $\sum_{n=1}^{+\infty} \alpha'(n) z^n$ with $\alpha'(n) = n\alpha(n)$. The function $\alpha'(n)$ satisfies the same assumptions as $\alpha(n)$ except with the abscissa $\sigma_0' = \sigma_0 + 1$ instead of $\sigma_0$. Therefore, using the homogeneity differential operator sufficiently many times, we can consider the function $n^k \alpha(n)$ with the abscissa $\sigma_0 + k$, where $k > -\sigma_0$. Applying the theorem for the abscissa $\sigma_0 + k > 0$, we deduce that $n^k \alpha(n) = n^{\sigma_0 + k - 1}$ for all $n$, that is, $\alpha(n) = n^{\sigma_0 - 1}$.

If $\sigma_0 > 0$, we can instead study $\tilde{\alpha}(n) = n^{\sigma_0 - 1} \alpha(n)$. The associated Dirichlet series $\widetilde{F}[\beta](s)$ satisfies $\widetilde{F}[\beta](s) = F[\beta](s + \sigma_0 - 1)$. If $\alpha(n)$ satisfies the assumptions of Theorem 18, then $\tilde{\alpha}(n)$ satisfies those of Theorem 1. We can deduce that there is a dense set of $\beta$ such that $\widetilde{F}[\beta](s)$ is analytical over $\Re s > \sigma_0$ and with $F[\beta](1+x) \to \infty$ as $x \to 0+$. By translation, we derive that for the same $\beta$, $F[\beta](s)$ is analytical over $\{\Re s > \sigma_0\}$ and with $F[\beta](\sigma_0 + x) \to +\infty$ as $x \to 0+$. By Proposition 16, $e(\beta)$ is a singular point of $\sum_{n=1}^{\infty} \alpha(n) z^n$ for a dense set of $\beta$, unless $\tilde{\alpha}(n) = 1$ for all $n$, that is, unless $\alpha(n) = n^{\sigma_0 - 1}$. □

## References

[Apostol 1976] T. M. Apostol, *Introduction to analytic number theory*, Springer, 1976. MR Zbl

[Hardy and Littlewood 1914] G. H. Hardy and J. E. Littlewood, "Tauberian theorems concerning power series and Dirichlet's series whose coefficients are positive", *Proc. London Math. Soc.* (2) **13**:1 (1914), 174–191. MR Zbl

[Lucht 1981] L. Lucht, "Power series with multiplicative coefficients", *Math. Z.* **177**:3 (1981), 359–374. MR Zbl

[Montgomery and Vaughan 2007] H. L. Montgomery and R. C. Vaughan, *Multiplicative number theory, I: Classical theory*, Cambridge Studies in Advanced Mathematics **97**, Cambridge University Press, 2007. MR Zbl

[Petrushov 2014] O. Petrushov, "On the behaviour close to the unit circle of the power series with Möbius function coefficients", *Acta Arith.* **164**:2 (2014), 119–136. MR Zbl

[Petrushov 2015a] O. Petrushov, "On the behavior close to the unit circle of the power series whose coefficients are squared Möbius function values", *Acta Arith.* **168**:1 (2015), 17–30. MR Zbl

[Petrushov 2015b] O. Petrushov, "On the behavior of power series with completely additive coefficients", *Bull. Pol. Acad. Sci. Math.* **63**:3 (2015), 217–225. MR Zbl

[Petrushov 2017] O. A. Petrushov, "On the behavior close to the unit circle of power series with additive coefficients", *Acta Arith.* **180**:4 (2017), 319–332. MR Zbl

[Petrushov 2018] O. A. Petrushov, "On the behavior of a power series with completely multiplicative coefficients near the unit circle", *Mat. Zametki* **103**:5 (2018), 750–764. In Russian; translated in *Math. Notes* **103**:5-6 (2018), 797–810. MR Zbl

Received 12 Sep 2018. Revised 21 Mar 2020.

OLEG A. PETRUSHOV:

olegap86@yandex.ru
APS RADIS, Moscow, Russia

msp

Moscow Journal of Combinatorics and Number Theory
Vol. 9, No. 2, 2020
https://doi.org/10.2140/moscow.2020.9.163

# On the roots of the Poupard and Kreweras polynomials

Frédéric Chapoton and Guo-Niu Han

The Poupard polynomials are polynomials in one variable with integer coefficients, with some close relationship to Bernoulli and tangent numbers. They also have a combinatorial interpretation. We prove that every Poupard polynomial has all its roots on the unit circle. We also obtain the same property for another sequence of polynomials introduced by Kreweras and related to Genocchi numbers. This is obtained through a general statement about some linear operators acting on palindromic polynomials.

## 1. Introduction

Let us consider the sequence of polynomials $(F_n)_{n\geq 1}$ in one variable $x$ characterized by the equation

$$(x-1)^2 F_{n+1}(x) = (x^{2n+2}+1)F_n(1) - 2x^2 F_n(x) \quad \text{for } n \geq 1, \tag{1-1}$$

with the initial condition $F_1 = 1$. When described in this way, their existence is not completely obvious, because the right-hand side must have a double root at $x = 1$ for the recurrence to make sense. The first few terms are given by

$$\begin{aligned} F_1 &= 1, \\ F_2 &= x^2 + 2x + 1, \\ F_3 &= 4x^4 + 8x^3 + 10x^2 + 8x + 4, \\ F_4 &= 34x^6 + 68x^5 + 94x^4 + 104x^3 + 94x^2 + 68x + 34. \end{aligned}$$

The polynomial $F_n$ has degree $2n-2$ and palindromic coefficients.

The coefficients of these polynomials form the Poupard triangle (A8301), first considered by Christiane Poupard [1989] and proved to enumerate some sets of labelled binary trees. It follows from this combinatorial interpretation that all coefficients of $F_n$ are nonnegative integers. For further combinatorial information on these polynomials and their relatives, see [Foata and Han 2013; 2014].

The constant terms of these polynomials form the sequence of *reduced tangent numbers* (A2105), which can be defined for $n \geq 1$ by the formula

$$\frac{2^n(2^{2n}-1)|B_{2n}|}{n}, \tag{1-2}$$

where $B_n$ are the classical Bernoulli numbers, and starts by

$$1,\ 1,\ 4,\ 34,\ 496,\ 11056,\ 349504,\ 14873104,\ 819786496,\ \ldots.$$

*MSC2010:* primary 26C10, 47B39; secondary 11B68, 39A70.
*Keywords:* palindromic polynomial, unit circle, complex root, linear operator, Bernoulli number.

One can deduce from (1-1) that $F_{n+1}(0) = F_n(1)$, so the reduced tangent numbers also describe the values of the polynomials $F_n$ at $x = 1$.

Our first result is the following unexpected property, which was the experimental starting point of this article.

**Theorem 1.1.** *For $n \geq 1$, all roots of the polynomial $F_n(x)$ are on the unit circle.*

This is proved in Section 2 in a much more general context, by showing that, for any positive integer $D$, a linear operator $\mathscr{N}_D$ maps palindromic polynomials with nonnegative coefficients to palindromic polynomials with nonnegative coefficients and all roots on the unit circle.

Whether there is any combinatorial meaning for this theorem, and for the similar theorem below, is rather unclear. Although the coefficients of these polynomials have a combinatorial interpretation, the location of their roots does not tell us anything about the combinatorics. One may speculate about some kind of arithmetic interpretation, maybe in terms of Weil polynomials, given the close relationship to Bernoulli numbers.

As another interesting application, one can consider the sequence of polynomials characterized by

$$(x-1)^2 G_{n+1}(x) = (x^{2n+3}+1)G_n(1) - 2x^2 G_n(x) \quad \text{for } n \geq 1, \tag{1-3}$$

with initial condition $G_1 = 1 + x$. The first few terms are

$$\begin{aligned} G_1 &= x + 1, \\ G_2 &= 2x^3 + 4x^2 + 4x + 2, \\ G_3 &= 12x^5 + 24x^4 + 32x^3 + 32x^2 + 24x + 12, \\ G_4 &= 136x^7 + 272x^6 + 384x^5 + 448x^4 + 448x^3 + 384x^2 + 272x + 136. \end{aligned}$$

The polynomial $G_n$ has degree $2n-1$ and palindromic coefficients.

**Theorem 1.2.** *For $n \geq 1$, all roots of the polynomial $G_n(x)$ are on the unit circle.*

Because the polynomials $G_n$ have odd degree, they are all divisible by $x + 1$. One can also show by induction that the polynomial $G_n$ is divisible by $2^{n-1}$. The quotient polynomials $2^{1-n}G_n/(x+1)$ have appeared in [Kreweras 1997], dealing with refined enumeration of some sets of permutations. Their constant terms are the Genocchi numbers (A1469), given by the formula

$$2(2^{2n} - 1)|B_{2n}|, \tag{1-4}$$

where $B_n$ are again the Bernoulli numbers.

Both theorems above are proved in Section 2 using a family of operators $\mathscr{N}_D$ acting on palindromic polynomials. Section 3 describes explicit simple eigenvectors of the operator $\mathscr{N}_1$. In Section 4, some evidence is given for the general asymptotic behaviour of the iteration of the operators $\mathscr{N}_D$ for $D > 1$. Section 5 contains various statements and conjectures on values of the operators $\mathscr{N}_D$ on specific palindromic polynomials.

Let us note as a side remark that another family of polynomials, also related to Bernoulli numbers, has been proved in [Lalín and Smyth 2013] to have only roots on the unit circle, by different methods.

## 2. Operators $\mathcal{N}_D$ and roots on the unit circle

Let us consider a polynomial $P(x) = \sum_{j=0}^{d} p_j x^j$ with rational coefficients. Let us say that the polynomial $P$ is *palindromic of index d* if $p_j = p_{d-j}$ for all $j$. Note that the index can also be described as the sum of the degree and the valuation. For example, the index of the polynomial $x = 0 + x + 0x^2$ is 2. For any $d \geq 0$, let $\mathsf{V}_d$ be the vector space spanned by palindromic polynomials of index $d$.

For every nonnegative integer $D$, let us introduce a linear operator $\mathcal{N}_D$ from $\mathsf{V}_d$ to $\mathsf{V}_{d+2D-2}$. This operator is characterized by the formula

$$(x-1)^2 \mathcal{N}_D(P)(x) = (x^{d+2D}+1)P(1) - 2x^D P(x). \tag{2-1}$$

The definition requires that the right-hand side is divisible by $(x-1)^2$. By the linearity of (2-1), it is enough to check this property for the basis elements $x^i + x^{d-i}$ with $0 \leq i \leq d$, where one finds

$$\mathcal{N}_D(x^i + x^{d-i}) = 2\frac{(1-x^{i+D})}{(1-x)}\frac{(1-x^{d+D-i})}{(1-x)}, \tag{2-2}$$

which is a polynomial with nonnegative integer coefficients. Note that when $d = 2i$, one can divide (2-2) by 2.

The definition of $\mathcal{N}_D$ and formula (2-2) imply immediately the following lemma.

**Lemma 2.1.** *Let $P$ be a nonzero palindromic polynomial of index $d$ with nonnegative integer coefficients. If $d \leq 1$, assume moreover that $D > 0$. Then $\mathcal{N}_D(P)$ is a nonzero palindromic polynomial of index $d + 2D - 2$ with positive integer coefficients.*

Let us record the following useful statement as a lemma.

**Lemma 2.2.** *When iterating $i$ times $\mathcal{N}_D$ on a palindromic polynomial $P$ of odd index with integer coefficients, the integer $2^i$ divides $\mathcal{N}_D^i P$.*

*Proof.* If the index of a palindromic polynomial $P$ is odd, then it is divisible by $x+1$. When $P$ has integer coefficients, (2-1) then implies that $\mathcal{N}_D(P)$ has one further factor 2. The lemma follows by induction. □

Recall that a palindromic polynomial $P = \sum_{j=0}^{d} p_j x^j$ is called *unimodal* if the sequence of coefficients is increasing up to the middle coefficient(s), then decreasing. A polynomial $P$ is called *concave* if the piecewise linear function that maps $j$ to $p_j$ is a concave function. A concave polynomial $P$ is called *strictly concave* if every point $(j, p_j)$ is moreover an extremal point in the graph of this piecewise linear function.

**Lemma 2.3.** *Let $P$ be a nonzero palindromic polynomial of index $d$ with nonnegative integer coefficients. If $d \leq 1$, assume moreover that $D > 0$. Then $\mathcal{N}_D(P)$ is unimodal and concave. If $P$ has no zero coefficient, then $\mathcal{N}_D(P)$ is strictly concave.*

*Proof.* By (2-2), the polynomial $\mathcal{N}_D(P)$ is a nonnegative linear combination of unimodal and concave polynomials, and hence is itself unimodal and concave. Each term in (2-2) gives two extremal points, or just one extremal point when $i = d - i$. When $P$ has no zero coefficient, this implies that there is an extremal point above every integer between 1 and $d+1$. □

Let us now recall a beautiful criterion obtained in [Lakatos and Losonczi 2004].

**Lemma 2.4.** *Let $P(x) = \sum_{j=0}^{d} p_j x^j$ be a palindromic polynomial of index $d$. If*

$$|p_d| \geq \frac{1}{2} \sum_{j=1}^{d-1} |p_j|, \tag{2-3}$$

*then all roots of $P$ are on the unit circle.*

The criterion above has been generalized recently in [Vieira 2017], which gives a sufficient condition for having a given number of roots on the unit circle.

From the criterion of Lemma 2.4, one deduces:

**Theorem 2.5.** *Let $P(x) = \sum_{j=0}^{d} p_j x^j$ be a palindromic polynomial of index $d$. If*

$$2p_j \geq p_{j-1} + p_{j+1} \quad \text{for all } 0 \leq j \leq d, \tag{2-4}$$

*with the convention that $p_{-1} = p_{d+1} = 0$, then all roots of $P$ are on the unit circle.*

*Proof.* Let $Q(x) = (1-x)^2 P(x)$. Then $Q(x) = \sum_{j=0}^{d+2} q_j x^j$, where

$$\begin{aligned} q_0 &= p_0, \\ q_{j+1} &= p_{j+1} + p_{j-1} - 2p_j \quad (0 \leq j \leq d), \\ q_{d+2} &= p_d. \end{aligned}$$

Note that $Q$ is also palindromic of index $d+2$.

By the hypothesis (2-4), all $q_j \leq 0$ for $1 \leq j \leq d+1$. Since $Q(1) = 0$, we have

$$\sum_{j=1}^{d+1} |q_j| = -\sum_{j=1}^{d+1} q_j = q_0 + q_{d+2} = 2q_{d+2}.$$

Note that therefore $q_0 \geq 0$.

Since $Q(x)$ is palindromic, and

$$|q_{d+2}| = \frac{1}{2} \sum_{j=1}^{d+1} |q_j|,$$

one can therefore apply Lemma 2.4 to $Q(x)$ and conclude that $Q(x)$ has all its roots on the unit circle. This implies the same property for $P(x)$. □

**Theorem 2.6.** *Let $P$ be a nonzero palindromic polynomial of index $d$ with nonnegative integer coefficients. If $d \leq 1$, assume moreover that $D > 0$. Then $\mathscr{N}_D(P)$ is a nonzero palindromic polynomial of index $d + 2D - 2$ with nonnegative integer coefficients, and all roots of $\mathscr{N}_D(P)$ are on the unit circle.*

*Proof.* This is an application of Theorem 2.5. The definition of $\mathscr{N}_D$ and the hypothesis that $P$ has nonnegative coefficients imply immediately the condition (2-4). □

Let us now apply Theorem 2.6 to the proofs of Theorems 1.1 and 1.2. The defining recurrence (1-1) for the polynomials $F_n$ can be written as $F_{n+1} = \mathscr{N}_D(F_n)$ with the initial condition $F_1 = 1$. The property follows by induction. The same proof works for $G_n$ with the initial polynomial $1 + x$.

Let us now state two useful lemmas.

**Lemma 2.7.** *For all $d \geq 0$, the polynomial $x^d + 1$ is in the kernel of $\mathscr{N}_0$.*

*Proof.* This is a direct consequence of (2-2). □

**Lemma 2.8.** *Let $d \geq 2$ be an integer. Then*

$$\mathscr{N}_0(1 + x + \cdots + x^d) = \sum_{i=0}^{d-2} (d-1-i)(i+1)x^i. \tag{2-5}$$

*Proof.* From the definition of $\mathscr{N}_0$ by (2-2), and by the previous lemma, this is equal to

$$\sum_{j=1}^{d-1} \frac{1-x^j}{1-x} \frac{1-x^{d-j}}{1-x}.$$

Expanding, one finds that the coefficient of $x^i$ is the cardinality of

$$\{(j,k) \mid 0 \leq k \leq j-1 \text{ and } 0 \leq i-k \leq d-j-1\}.$$

But this is the same as the set

$$\{(j,k) \mid 1 \leq j-k \leq d-i-1 \text{ and } 0 \leq k \leq i\},$$

whose cardinality is $(d-1-i)(i+1)$. □

## 3. Sinus polynomials as eigenvectors

As can be seen in Figure 1, right, the roots of the Poupard polynomials $F_n(x)$ are very close to some of the roots of $x^{2n}+1$, with two missing roots on the right. Moreover the plot of the coefficients of $F_n(x)$ seems to approximate a concave continuous function, as in Figure 1, left.

One expects that, up to a global multiplicative factor, the polynomials obtained when iterating $n$ times the operator $\mathscr{N}_D$ (for some fixed $D > 1$) are always becoming, when $n$ is large, very close to the polynomials described in this section. Some kind of justification will be given in the next section.

Let us consider the polynomial $S_{m,n}(x)$ defined for $n \geq 2$ and odd $m \geq 1$ by

$$S_{m,n}(x) = \frac{x^{mn}+1}{x^2 - 2x\cos\frac{\pi}{n} + 1}, \tag{3-1}$$

whose roots are the roots of $x^{mn}+1$ except $\exp(i\pi/n)$ and its conjugate.

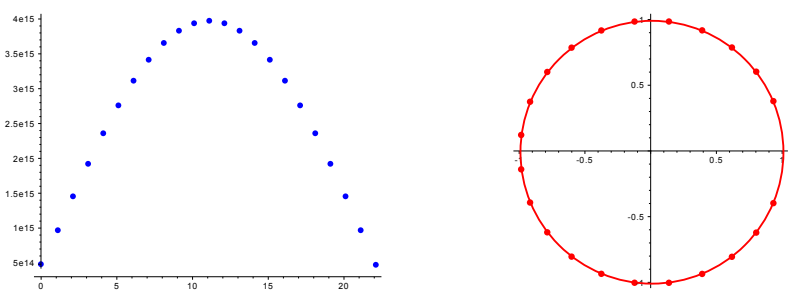



**Figure 1.** Coefficients and roots of the Poupard polynomial $F_{12}$.

Let us first give an alternative expression for $S_{m,n}$.

**Lemma 3.1.** *The polynomial $S_{m,n}$ has the explicit expression*

$$S_{m,n}(x) = \frac{1}{\sin\frac{\pi}{n}} \sum_{k=0}^{mn-2} \sin\left(\frac{(k+1)\pi}{n}\right) x^k. \tag{3-2}$$

*Proof.* The proof is a simple computation, expanding both sides as polynomials in $x$ and $\zeta = \exp(i\pi/n)$, also using that $m$ is odd. □

This implies that the plot of the coefficients of $S_{1,n}$ looks very much like a sinus curve, like Figure 1, left.

**Proposition 3.2.** *For every $n \geq 2$ and odd $m \geq 1$, the polynomial $S_{m,n}$ is an eigenvector of the operator $\mathscr{N}_1$ acting on $\mathsf{V}_{mn-2}$ for the eigenvalue $1/(1-\cos(\pi/n))$.*

*Proof.* The proof is another explicit computation using the definition of $S_{m,n}$ in (3-1) and the definition of the operator $\mathscr{N}_1$ in (2-1). □

Note that the eigenvalue is also the value $S_{m,n}(1)$.

In general, the Galois conjugates of the polynomial $S_{1,n}$ do not provide a complete set of eigenvectors for the operator $\mathscr{N}_1$ acting on $\mathsf{V}_{n-2}$. The other eigenvectors are $S_{m,n/m}$ for odd divisors $m$ of $n$, and their Galois conjugates.

The family of operators $\mathscr{N}_1$ acting on the spaces $\mathsf{V}_{n-2}$ of palindromic polynomials looks very much like discrete versions of the Laplacian operator $\partial_x^2$ acting on the space of functions $f$ on the real interval $[0, 1]$ such that $f(1-x) = f(x)$ for all $x$ and $f(0) = f(1) = 0$.

## 4. Asymptotic behaviour from recurrence

Our next point is to justify in a heuristic way that iterating an operator $\mathscr{N}_D$ for some $D > 1$ produces a sequence of polynomials that gets closer and closer to the sinus polynomials $S_{1,n}$. We have not tried to make these computations rigorous.

Let us consider a family of polynomials $H_n$ of index $n$ defined by iterating $\mathscr{N}_D$, starting from an arbitrary palindromic polynomial $H_m$ with nonnegative coefficients and index $m$. Throughout this section, the index $n$ belongs to an arithmetic progression of step $\delta = 2D-2$ starting at $m$. Let us write

$$H_n(x) = \sum_{k=0}^{n} H_{n,k} x^k. \tag{4-1}$$

We will assume the following asymptotic ansatz for the constant terms:

$$H_n(0) \simeq \mathtt{A}\mathtt{B}^n n^{\mathtt{C}} n^{\mathtt{E}n} \tag{4-2}$$

for some constants $\mathtt{A}, \mathtt{B}, \mathtt{C}, \mathtt{E}$, with $\mathtt{A}, \mathtt{B}, \mathtt{E}$ positive. This ansatz is motivated by the known case of the tangent numbers, where $\mathtt{B} = 2/(e\pi)$, $\mathtt{E} = 1$ and $\mathtt{C} = -\frac{1}{2}$. This ansatz implies

$$H_{n+\delta}(0)/H_n(0) \simeq \mathtt{B}^{\delta} e^{\delta\mathtt{E}} n^{\delta\mathtt{E}}. \tag{4-3}$$

We will also assume that there exists a smooth function $\Psi$ which is a probability distribution function on the real interval [0, 1] vanishing at 0 and 1, with $\Psi(1-x)=\Psi(x)$ on this interval and such that

$$H_{n,k} \simeq \frac{\alpha_n}{n}\Psi\left(\frac{k}{n}\right) + H_{n,0} \tag{4-4}$$

is a good asymptotic approximation when $n$ is large, for some sequence $\alpha_n$ to be determined.

Taking the sum of (4-4) over $k$ ranging from 0 to $n$ and using the hypothesis on $\Psi$, one gets

$$H_{n+\delta,0} = H_n(1) \simeq \alpha_n + (n+1)H_{n,0}.$$

Assuming that $nH_{n,0}$ is negligible compared to $H_{n+\delta,0}$, one obtains that a correct choice for $\alpha_n$ is

$$\alpha_n = H_{n+\delta,0}.$$

From (2-1), one deduces that the action of $\mathscr{N}_D$ at the level of coefficients is given by

$$H_{n+\delta,k+D} - 2H_{n+\delta,k+D-1} + H_{n+\delta,k+D-2} = -2H_{n,k}, \tag{4-5}$$

except for $k=0$ and $k=n$.

Replacing in (4-5) the coefficients by the expression from (4-4), one obtains

$$\frac{\alpha_{n+\delta}}{n+\delta}\left(\Psi\left(\frac{k+D}{n+\delta}\right) - 2\Psi\left(\frac{k+D-1}{n+\delta}\right) + \Psi\left(\frac{k+D-2}{n+\delta}\right)\right) \simeq -2\left(\frac{\alpha_n}{n}\Psi\left(\frac{k}{n}\right) + H_{n,0}\right). \tag{4-6}$$

Using now the growth ansatz, one can get rid of $H_{n,0}$ in the rightmost term and obtain

$$\Psi\left(\frac{k+D}{n+\delta}\right) - 2\Psi\left(\frac{k+D-1}{n+\delta}\right) + \Psi\left(\frac{k+D-2}{n+\delta}\right) \simeq -2\frac{\alpha_n}{\alpha_{n+\delta}}\Psi\left(\frac{k}{n}\right). \tag{4-7}$$

The left-hand side is an approximation of the second derivative of $\Psi$, so one obtains

$$\frac{1}{2(n+\delta)^2}\Psi''\left(\frac{k}{n+\delta}\right) \simeq -2\frac{\alpha_n}{\alpha_{n+\delta}}\Psi\left(\frac{k}{n}\right). \tag{4-8}$$

If $\delta\mathtt{E}=2$, one therefore reaches the differential equation

$$\Psi'' = -\mathtt{F}\Psi, \tag{4-9}$$

where $\mathtt{F} = 4/(\mathtt{B}^\delta e^2)$. Because $\Psi$ vanishes at 0, it must be a multiple of $\sin(\sqrt{\mathtt{F}}x)$. Because $\Psi$ vanishes at 1 and is positive on the interval [0, 1], necessarily $\mathtt{F}=\pi^2$ and therefore $\mathtt{B}^\delta = (2/(e\pi))^2$. Because $\Psi$ is a probability distribution, one must have $\Psi = \frac{\pi}{2}\sin(\pi x)$.

One can therefore conclude that, under several plausible but unproven assumptions, the asymptotic shape of the coefficients of the polynomials $H_n$ is approximating that of the polynomials $S_{1,n+2}$.

## 5. Various remarks

**5.1.** ***Action of the operator*** $\mathscr{N}_0$**.** Applying the operator $\mathscr{N}_0$ decreases the index by 2, so that iterating this operator on any initial polynomial $P$ of index $d$ always vanishes after a finite number of steps. Let $\mathscr{N}_0^{\max}$ be the last nonidentically zero iterate of $\mathscr{N}_0$ acting on $\mathsf{V}_d$. Let us denote by $\rho$ the linear map that maps $P$ to the constant term of $\mathscr{N}_0^{\max}(P)$.

For example, here is a sequence of iterates of $\mathscr{N}_0$:

$$x^4+x^3+x^2+x+1, \quad 3x^2+4x+3, \quad 4.$$

In this case, $\rho(x^4+x^3+x^2+x+1)=4$.

Let us present some special cases of initial choices where the value of $\rho$ is interesting.

For $n \geq 0$, consider the polynomial

$$Q_n(t) = \sum_{i=0}^{2n+1} \rho\left(\frac{x^i - x^{2n+1-i}}{x-1}\right) t^i, \tag{5-1}$$

recording this sequence of final values. By the antisymmetry of the argument of $\rho$, the polynomial $Q_n$ vanishes at $t=1$. Let $P_n(t)$ be the quotient $Q_n(t)/(t-1)$, which is clearly a palindromic polynomial.

**Proposition 5.1.** *For every $n \geq 0$, the polynomial $P_n$ is the Poupard polynomial $F_{n+1}$.*

*Proof.* For $n=0$, one can check that $P_n(t)=1$. Assume $n>0$. For $0 \leq i \leq 2n$, the coefficient $c_{n,i}$ of $t^i$ in $P_n(t)$ can be written as

$$-\rho\left(\sum_{0\leq k\leq i} \frac{x^k - x^{2n+1-k}}{x-1}\right) = \rho\left(\frac{x^{i+1}-1}{x-1}\frac{x^{2n+1-i}-1}{x-1}\right). \tag{5-2}$$

Let us now compute $c_{n,i+2} - 2c_{n,i+1} + c_{n,i}$ for $0 \leq i \leq 2n-2$. Starting from the left-hand side of (5-2), this is given by

$$\rho(x^{i+1} + x^{2n-i-1}).$$

Using now (2-2) for $\mathscr{N}_0$ and the definition of $\rho$ as the final value for the iteration of $\mathscr{N}_0$, this becomes

$$2\rho\left(\frac{x^{i+1}-1}{x-1}\frac{x^{2n-i-1}-1}{x-1}\right),$$

in which one can recognize $-2c_{n-1,i}$ using the right-hand side of (5-2).

Moreover, $c_{n,1} - 2c_{n,0} = \rho(1+x^{2n}) = 0$ because $1+x^{2n}$ is in the kernel of $\mathscr{N}_0$ by Lemma 2.7.

Let us now check that $c_{n,0} = \sum_{i=0}^{2n-2} c_{n-1,i}$. First, by (5-2), the left-hand side is the image by $\rho$ of $\mathscr{N}_0(1+x+\cdots+x^{2n})$, given by Lemma 2.8. The right-hand side is the image by $\rho$ of

$$\sum_{i=0}^{2n-2} \sum_{0\leq k\leq i} \sum_{k\leq j\leq 2n-2-k} x^j = \sum_{j=0}^{2n-2} (2n-1-j)(j+1)x^j, \tag{5-3}$$

which is the exact same expression.

All these properties of the coefficients $c_{n,i}$ imply exactly that the polynomial $P_n(t)$ is the image of $P_{n-1}(t)$ by $\mathscr{N}_1$, acting on the variable $t$. □

For $n \geq 0$, consider the polynomial

$$Q'_n(t) = \sum_{i=0}^{2n} \rho\left(\frac{x^i - x^{2n-i}}{x-1}\right) t^i, \tag{5-4}$$

recording this sequence of final values. By the antisymmetry of the argument of $\rho$, the polynomial $Q'_n$ vanishes at $t=1$. Let $P'_n(t)$ be the quotient $Q'_n(t)/(t-1)$, which is clearly a palindromic polynomial of odd index.

**Proposition 5.2.** *For every $n \geq 1$, the polynomial $Q'_n$ is the Kreweras polynomial $G_n$.*

*Proof.* The proof is very similar to the previous one. One first check that $Q'_1$ is $1+x$. Then one checks by looking at coefficients that $Q'_{n+1}$ is $\mathscr{N}_1(Q'_n)$. □

Let us now describe some similar conjectural properties. For the starting sequence $(2^{-j}(1+x)^{2j})_{j\geq 0}$, one gets the following values of $\rho$:

$$1,\ 1,\ 5,\ 61,\ 1385,\ 50521,\ 2702765,\ 199360981,\ 19391512145,\ \ldots,$$

which seem to be the Euler numbers (A364). Similarly, for the starting sequence $(2^{-j}(1+x)^{2j+1})_{j\geq 0}$, one gets

$$1,\ 3,\ 25,\ 427,\ 12465,\ 555731,\ 35135945,\ \ldots.$$

This seems to be the closely related sequence (A9843).

As a final conjectural remark, let us consider the following extension of the two previous cases.

**Conjecture 5.3.** *For every $i$, $j$, the number $\rho(x^i(1+x)^j)$ is divisible by $2^{\lfloor j/2\rfloor}$.*

This property is clear if $j$ is odd by Lemma 2.2, but not at all if $j$ is even.

Assuming this conjecture, one can define, for every integer $n$, the square matrix $\boldsymbol{M}_n$ whose coefficient $\boldsymbol{M}_n(i,j)$, for $0 \leq i \leq n$ and $0 \leq j \leq n$, is $\rho(x^i(1+x)^j)2^{-\lfloor j/2\rfloor}$.

**Conjecture 5.4.** *For all $n \geq 0$, the determinant $d_n$ of the matrix $\boldsymbol{M}_n$ is given by the formula*

$$d_n = (n-1)!^{\varepsilon(1)}(n-2)!^{\varepsilon(2)}(n-3)!^{\varepsilon(3)}\cdots 1!^{\varepsilon(n-1)}, \tag{5-5}$$

*where*

$$\varepsilon(k) = \begin{cases} 2 & \text{if } k \text{ is odd,} \\ 4 & \text{if } k \text{ is even.} \end{cases}$$

For example, $\boldsymbol{M}_6$ is equal to

$$\begin{pmatrix} 1 & 1 & 1 & 3 & 5 & 25 \\ 1 & 2 & 3 & 14 & 33 & 226 \\ 2 & 8 & 18 & 120 & 378 & 3336 \\ 10 & 64 & 198 & 1728 & 6858 & 74304 \\ 104 & 896 & 3528 & 38016 & 182088 & 2339712 \\ 1816 & 19456 & 92808 & 1188864 & 6668568 & 99118080 \end{pmatrix}, \tag{5-6}$$

whose determinant is indeed $5!^2\,4!^4\,3!^2\,2!^4\,1!^2$.

This matrix contains entries with large prime factors, for example $92808 = 2^3 3^2 1289$, but the determinant has only small prime factors.

**5.2.** ***Action of the operator $\mathscr{N}_1$.*** Applying the operator $\mathscr{N}_1$ does not change the index, so iterating this operator on any initial choice gives an infinite sequence of palindromic polynomials of the same index.

For example, starting with $x$ gives a sequence of polynomials

$$x,\quad x^2+2x+1,\quad 4x^2+6x+4,\quad 14x^2+20x+14,\quad 48x^2+68x+48,\quad 164x^2+232x+164,\quad \ldots,$$

whose constant terms and middle coefficients are given by A7070 and by A6012. Indeed, the action of $\mathscr{N}_1$ on reciprocal polynomials of index 2 is given in the basis $\{1+x^2, x\}$ by the matrix

$$\begin{pmatrix} 2 & 1 \\ 2 & 2 \end{pmatrix}$$

so that both sequences satisfy the recurrence $a_n = 4a_{n-1} - 2a_{n-2}$ with appropriate initial conditions.

**5.3.** ***Action of the operator*** $\mathscr{N}_2$**.** Applying the operator $\mathscr{N}_2$ increases the index by 2, so iterating this operator gives an infinite sequence of polynomials for every initial choice. In each such sequence, the sequence of constant terms is, up to a shift of indices by 1, the same as the sequence of values at $x = 1$. Some examples were presented in the introduction, related to reduced tangent numbers and Genocchi numbers. Let us record one more family of examples.

Using the polynomials $x^i(x+1)$ for $i \geq 0$ as starting points, one gets a table of constant terms:

$$\begin{pmatrix} 1 & 1 & 3 & 17 & 155 & 2073 \\ 0 & 1 & 6 & 55 & 736 & 13573 \\ 0 & 1 & 10 & 135 & 2492 & 60605 \\ 0 & 1 & 15 & 280 & 6818 & 211419 \\ 0 & 1 & 21 & 518 & 16086 & 619455 \\ 0 & 1 & 28 & 882 & 34020 & 1592811 \end{pmatrix}.$$

Here $i$ is the row index and in each row the term of index $j$ is the constant term divided by $2^j$. This table seems to be essentially the Salié triangle (A65547).

## References

[Foata and Han 2013] D. Foata and G.-N. Han, "Finite difference calculus for alternating permutations", *J. Difference Equ. Appl.* **19**:12 (2013), 1952–1966. MR Zbl

[Foata and Han 2014] D. Foata and G.-N. Han, "Tree calculus for bivariate difference equations", *J. Difference Equ. Appl.* **20**:11 (2014), 1453–1488. MR Zbl

[Kreweras 1997] G. Kreweras, "Sur les permutations comptées par les nombres de Genocchi de 1-ière et 2-ième espèce", *European J. Combin.* **18**:1 (1997), 49–58. MR Zbl

[Lakatos and Losonczi 2004] P. Lakatos and L. Losonczi, "Self-inversive polynomials whose zeros are on the unit circle", *Publ. Math. Debrecen* **65**:3-4 (2004), 409–420. MR Zbl

[Lalín and Smyth 2013] M. N. Lalín and C. J. Smyth, "Unimodularity of zeros of self-inversive polynomials", *Acta Math. Hungar.* **138**:1-2 (2013), 85–101. Addendum in **147**:1 (2015), 255–257. MR Zbl

[Poupard 1989] C. Poupard, "Deux propriétés des arbres binaires ordonnés stricts", *European J. Combin.* **10**:4 (1989), 369–374. MR Zbl

[Vieira 2017] R. S. Vieira, "On the number of roots of self-inversive polynomials on the complex unit circle", *Ramanujan J.* **42**:2 (2017), 363–369. MR Zbl

Received 16 Jan 2020. Revised 15 May 2020.

FRÉDÉRIC CHAPOTON:
chapoton@unistra.fr
Institut de Recherche Mathématique Avancée, UMR 7501, Université de Strasbourg et CNRS, Strasbourg, France

GUO-NIU HAN:
guoniu.han@unistra.fr
Institut de Recherche Mathématique Avancée, UMR 7501, Université de Strasbourg et CNRS, Strasbourg, France

msp

Moscow Journal of Combinatorics and Number Theory
Vol. 9, No. 2, 2020
https://doi.org/10.2140/moscow.2020.9.173

# Generalized colored circular palindromic compositions

Petros Hadjicostas

We derive the generating function (g.f.) of the number of colored circular palindromic compositions of $N$ with $K$ parts in terms of the g.f. of an input sequence $a$ that determines how many different colors each part of the composition can have. As a result, we get the g.f. of the number of all colored circular palindromic compositions of $N$. Using the latter formula and the g.f. of the number of colored circular compositions, we may easily derive the g.f. of the number of all colored dihedral compositions of $N$.

## 1. Introduction

A *linear composition* of a positive integer $N$ with length $K$ is a $K$-tuple $(\lambda_1, \dots, \lambda_K) \in \mathbb{Z}_{>0}^K$ such that

$$N = \lambda_1 + \cdots + \lambda_K.$$

Here the numbers $\lambda_1, \dots, \lambda_K$ are called *parts* of the composition. For example, $(1, 2, 3, 3)$ and $(3, 3, 2, 1)$ are two different linear compositions of $N = 9$ with $K = 4$ parts each.

*Cyclic or circular compositions* of $N$ with length $K$ are equivalence classes on the set of all linear compositions of $N$ with length $K$ such that two compositions are equivalent if and only if one can be obtained from the other by a cyclic shift. For example, $\{(1, 2, 3, 3), (3, 1, 2, 3), (3, 3, 1, 2), (2, 3, 3, 1)\}$ and $\{(3, 3, 2, 1), (1, 3, 3, 2), (2, 1, 3, 3), (3, 2, 1, 3)\}$ are two different cyclic or circular compositions of $N = 9$ with $K = 4$ parts each.

Cyclic or circular compositions were studied, for example, in [Ferrari and Zagaglia Salvi 2018; Gibson et al. 2018; Hadjicostas 2016; 2017; Knopfmacher and Robbins 2010; Sommerville 1909; Zagaglia Salvi 1999].

A (*bilaterally*) *symmetric cyclic composition* is a *circular palindrome*, that is, a circular or cyclic composition with at least one axis of (reflective) symmetry. Such circular or cyclic palindromic compositions were studied in [Bower 2010; Hadjicostas and Zhang 2017; Sommerville 1909; Williamson 1972].

For example, $\{(1, 3, 2, 3), (3, 1, 3, 2), (2, 3, 1, 3), (3, 2, 3, 1)\}$ is a circular palindromic composition of $N = 9$ with $K = 4$ parts with one axis of symmetry (through 1 and 2). On the other hand, $\{(1, 2, 1, 2), (2, 1, 2, 1)\}$ is a circular palindromic composition of $N = 6$ with $K = 4$ parts and two axes of symmetry (through the two 1's and through the two 2's). Finally, $\{(1, 1, 2, 2), (2, 1, 1, 2), (2, 2, 1, 1), (1, 2, 2, 1)\}$ is a circular palindromic composition of $N = 6$ with $K = 4$ parts and one axis of symmetry (that passes through no part).

*Dihedral compositions* of $N$ with length $K$ are equivalence classes on the set of all linear compositions of $N$ with length $K$ such that two compositions are equivalent if and only if one can be obtained from the

*MSC2010:* 05A10, 05A15, 11B39.
*Keywords:* circular palindrome, dihedral composition, generating function.

other by a cyclic shift or a reversal of order. Such compositions were studied, for example, in [Hadjicostas 2017; Knopfmacher and Robbins 2013; Zagaglia Salvi 1999]. For example,

$$\{(1,2,3,3),(3,1,2,3),(3,3,1,2),(2,3,3,1),(3,3,2,1),(1,3,3,2),(2,1,3,3),(3,2,1,3)\}$$

is a dihedral composition of $N=9$ with $K=4$ parts.

A *colored linear composition* of $N$ with $K$ parts according to an *input sequence*

$$a=(a(m):m\in\mathbb{Z}_{>0}),\quad \text{where } a(m)\in\mathbb{Z}_{\geq 0} \text{ for each } m\in\mathbb{Z}_{>0},$$

is a $K$-tuple $((\lambda_1,m_1),\dots,(\lambda_K,m_K))$ such that $\lambda_i,m_i\in\mathbb{Z}_{>0}$ for $i=1,\dots,K$ with

$$N=\lambda_1+\cdots+\lambda_K \quad\text{and}\quad 1\leq m_i\leq a(\lambda_i) \quad \text{for } i=1,\dots,K. \tag{1-1}$$

Note that, when $a(m)=0$ for some $m\in\mathbb{Z}_{>0}$, no part $\lambda_i$ of the colored linear compositions of $N$ we are considering (according to the input sequence $a$) can equal $m$.

For example, $(1_2,2_3,1_3,2_1)$ is a colored linear composition of $N=6$ with $K=4$ parts, where the parts have colors 2, 3, 3 and 1, respectively. This is a colored composition with respect to any input sequence $a=(a(m):m\in\mathbb{Z}_{>0})$ that satisfies $a(m)\in\mathbb{Z}_{\geq 0}$ for each $m\in\mathbb{Z}_{>0}$ and the inequalities $a(1)\geq 2$ and $a(2)\geq 2$.

*Colored cyclic compositions*, *colored* (*bilaterally*) *symmetric cyclic compositions* (i.e., *colored palindromic cyclic compositions*), and *colored dihedral compositions* of $N$ with length $K$ can be defined much as above. Colored compositions (of any kind) were studied in [Agarwal 2000; 2003; Bower 2010; Heubach and Mansour 2010, Section 3.5; Gibson et al. 2018].

For example, $\{(1_2,2_3,1_3,2_1),(2_3,1_3,2_1,1_2),(1_3,2_1,1_2,2_3),(2_1,1_2,2_3,1_3)\}$ is a circular nonpalindromic (not bilaterally symmetric) colored composition of $N=6$ with $K=4$ parts. On the other hand, $\{(1_2,2_3,1_3,2_3),(2_3,1_3,2_3,1_2),(1_3,2_3,1_2,2_3),(2_3,1_2,2_3,1_3)\}$ is a circular palindromic colored composition of $N=6$ with $K=4$ parts and one axis of symmetry (through $1_2$ and $1_3$).

Consider the collection $\mathcal{C}(N,K;a)$ of all colored compositions of $N$ with $K$ parts according to an *input sequence* $a=(a(m):m\in\mathbb{Z}_{>0})$. Consider a partition $\mathcal{P}=\mathcal{P}(N,K;a)$ of $\mathcal{C}(N,K;a)$ into (nonempty) *equivalence classes*. Denote by $b^{\mathcal{P}}(N,K)=b^{\mathcal{P}}(N,K;a)$ the total number of such equivalence classes in $\mathcal{P}$. Throughout the paper, instead of using $\mathcal{P}$, we use different superscripts (e.g., L, PL, CP, etc.) to denote the partition $\mathcal{P}$ of $\mathcal{C}(N,K;a)$. In particular, when each equivalence class in $\mathcal{P}$ has only one element, we use the superscript L (which stands for *linear* compositions).

Given $K\in\mathbb{Z}_{>0}$ and an input sequence $a=(a(m):m\in\mathbb{Z}_{>0})$, consider a sequence of partitions $\mathcal{P}_{K,a}=(\mathcal{P}(N,K;a):N\in\mathbb{Z}_{>0})$ and the set

$$\bigcup_{N\in\mathbb{Z}_{>0}}\mathcal{C}(N,K;a)$$

of all colored compositions of positive integers with $K$ parts such that $\mathcal{P}(N,K;a)$ is a partition of $\mathcal{C}(N,K;a)$ for each $N\in\mathbb{Z}_{>0}$. We call the sequence

$$b_K=(b(N,K):N\in\mathbb{Z}_{>0})$$

the corresponding *output sequence*, where for simplicity we have dropped the superscript $\mathcal{P}_{K,a}$ from $b_K$ and the superscript $\mathcal{P}(N,K;a)$ from $b(N,K)$.

Given $a = (a(m) : m \in \mathbb{Z}_{>0})$, we may also consider a family of partitions

$$\mathcal{P}_a = (\mathcal{P}(N, K; a) : N, K \in \mathbb{Z}_{>0} \text{ with } 1 \le K \le N)$$

and the set

$$\bigcup_{\substack{N,K\in\mathbb{Z}_{>0}\\ 1\le K\le N}} \mathcal{C}(N, K; a)$$

of all colored compositions with input sequence $a$ such that $\mathcal{P}(N, K; a)$ is a partition of $\mathcal{C}(N, K; a)$ for each pair $(N, K)$. The $N$-th term of the sequence

$$b = (b(N) : N \in \mathbb{Z}_{>0}) = \left(\sum_{K=1}^{N} b(N, K) : N \in \mathbb{Z}_{>0}\right)$$

gives the *total* number of equivalence classes of colored compositions of $N$ according to the family of partitions $(\mathcal{P}(N, K; a) : 1 \le K \le N)$. Again, for simplicity, we have dropped the superscript $\mathcal{P}_a$ from $b$ and $b(N)$ and the superscript $\mathcal{P}(N, K; a)$ from $b(N, K)$.

Throughout the paper, for integers $N$ and $K$, we set (for convenience)

$$b(N, K) = 0 \quad \text{when } K < 1 \text{ or } K > N. \tag{1-2}$$

Following [Bower 2010], we denote the (formal) *generating functions* (g.f.'s) of the three sequences $a$, $b_K$, and $b$ by

$$A(x) = \sum_{m=1}^{\infty} a(m)\, x^m, \quad B_K(x) = \sum_{N=1}^{\infty} b(N, K)\, x^N,$$
$$B(x) = \sum_{N=1}^{\infty} b(N)\, x^N = \sum_{N=1}^{\infty}\sum_{K=1}^{N} b(N, K)\, x^N,$$

respectively. The following trivial result connects $B_K(x)$ with $B(x)$ under mild assumptions.

**Proposition 1.1.** *If $b(N, K) = 0$ for all $K \ge 2$ and $N \in \{1, \dots, K-1\}$, then*

$$B(x) = \sum_{K=1}^{\infty} B_K(x).$$

As mentioned above, we use the superscript L to denote linear colored compositions of $N$ according to some input sequence $a = (a_m : m \in \mathbb{Z}_{>0})$ with g.f. $A(x)$. It is well known that

$$\begin{aligned} B_K^{\mathrm{L}}(x) &= \sum_{N=1}^{\infty} b^{\mathrm{L}}(N, K)\, x^N = A(x)^K \quad \text{for } K \in \mathbb{Z}_{>0},\\ B^{\mathrm{L}}(x) &= \sum_{N=1}^{\infty} b^{\mathrm{L}}(N)\, x^N = \frac{A(x)}{1 - A(x)}. \end{aligned} \tag{1-3}$$

This is the INVERT transform in [Bernstein and Sloane 1995, p. 61] and the AIK transform in [Bower 2010].

**Example 1.2.** When the input sequence is $a = (m : m \in \mathbb{Z})$, then we are dealing with the so-called *m-colored* (*linear*) *compositions* of $N$ studied in [Abrate et al. 2014; Agarwal 2000; Heubach and Mansour 2010, Section 3.5]. In such a case, $A(x) = x/(1-x)^2$ with

$$B_K^{\mathrm{L}}(x) = \frac{x^K}{(1-x)^{2K}} \quad \text{for } K \in \mathbb{Z}_{>0} \qquad \text{and} \qquad B^{\mathrm{L}}(x) = \frac{x}{x^2-3x+1}. \tag{1-4}$$

From this, we may easily deduce that

$$b^{\mathrm{L}}(N, K) = \binom{N+K-1}{2K-1} \quad \text{for } K \in \mathbb{Z}_{>0} \qquad \text{and} \qquad b^{\mathrm{L}}(N) = F_{2N}, \tag{1-5}$$

where $F_n$ is the $n$-th *Fibonacci number*. Other properties of "$n$-color compositions" can be found in [Agarwal 2003; Guo 2012; Sachdeva and Agarwal 2017].

The organization of the paper is as follows. In Section 2, we study colored linear palindromic compositions and derive their generating functions (see Theorems 2.1 and 2.2). The material in that section is needed for the material in Section 3, which is the main section of the paper. In Section 3, we study colored circular palindromic compositions and derive their generating functions (see Theorems 3.1 and 3.2, which generalize results in [Hadjicostas and Zhang 2017]). In particular, we prove that, if $b^{\mathrm{CP}}(N)$ is the total number of colored circular palindromic compositions of $N$ with input sequence $a = (a(m) : m \in \mathbb{Z}_{>0})$, then the g.f. of the sequence $(b^{\mathrm{CP}}(N) : N \in \mathbb{Z}_{>0})$ is given by

$$\sum_{N=1}^{\infty} b^{\mathrm{CP}}(N)\, x^N = \frac{(1+A(x))^2}{2(1-A(x^2))} - \frac{1}{2}, \tag{1-6}$$

where $A(x) = \sum_{m=1}^{\infty} a(m)\, x^m$.

Equation (1-6) is important because it allows the calculation of the g.f. of the number of colored dihedral compositions of $N$, say $b^D(N)$. When $A(x)$ is the g.f. of the input sequence $a = (a(m) : m \in \mathbb{Z}_{>0})$, then it can be proved (using Möbius inversion) that the g.f. of the number of colored circular compositions of $N$ is $-\sum_{d=1}^{\infty} (\phi(d)/d) \log(1 - A(x^d))$, where $\phi(d)$ is the Euler totient function. We omit the details of the proof of this result, but see, for example, [Flajolet and Sedgewick 2009; Flajolet and Soria 1991]. The g.f. of the number of colored dihedral compositions is then given by

$$\sum_{N=1}^{\infty} b^D(N)\, x^N = -\frac{1}{2} \sum_{d=1}^{\infty} \frac{\phi(d)}{d} \log(1 - A(x^d)) + \frac{(1+A(x))^2}{4(1-A(x^2))} - \frac{1}{4}. \tag{1-7}$$

## 2. Colored linear palindromic compositions

The study of *linear palindromic compositions* goes back all the way to the 19th century work [MacMahon 1893]. These compositions have been studied by several mathematicians since then. They are compositions $(\lambda_1, \ldots, \lambda_K)$ of $N$ with $K$ parts such that $\lambda_{K+1-i} = \lambda_i$ for $i = 1, \ldots, K$. Hadjicostas and Zhang [2017] called these compositions *type-I palindromic compositions* and denoted their number by $P^{L_1}(N, K)$. MacMahon [1893] proved that, for $n, k \in \mathbb{Z}_{>0}$ with $1 \le k \le n$,

$$P^{L_1}(2n, 2k) = P^{L_1}(2n, 2k-1) = P^{L_1}(2n-1, 2k-1) = \binom{n-1}{k-1},$$

while $P^{L_1}(2n-1, 2k) = 0$. We shall not use the notation $P^{L_1}$ in this paper.

Using the superscript PL for colored linear palindromic compositions and the superscript L for (general) linear compositions of $N$ with $K$ parts according to the input sequence $a = (a(m) : m \in \mathbb{Z}_{>0})$ in both cases, it is easy to prove (e.g., see the LPAL transform in [Bower 2010]) that:

- If $N$ and $K$ are even, then $b^{\mathrm{PL}}(N, K) = b^{\mathrm{L}}(N/2, K/2)$.
- If $N$ is odd and $K$ is even, then $b^{\mathrm{PL}}(N, K) = 0$.
- If $K = 1$, then $b^{\mathrm{PL}}(N, K = 1) = a(N)$.
- If $N$ is even and $K$ is odd $\geq 3$, then

$$b^{\mathrm{PL}}(N, K) = \sum_{0<i<N/2} a(2i)\, b^{\mathrm{L}}\Big(\frac{N}{2} - i, \frac{K-1}{2}\Big).$$

- If $N$ and $K$ are both odd with $K \geq 3$, then

$$b^{\mathrm{PL}}(N, K) = \sum_{0<i<N/2} a(2i-1)\, b^{\mathrm{L}}\Big(\frac{N+1}{2} - i, \frac{K-1}{2}\Big).$$

**Theorem 2.1.** *For fixed $K \in \mathbb{Z}_{>0}$, the g.f. of the sequence $(b^{\mathrm{PL}}(N, K) : N \in \mathbb{Z}_{>0})$ is given by*

$$B_K^{\mathrm{PL}}(x) = \begin{cases} A(x^2)^{K/2} & \text{if } K \text{ is even,} \\ A(x)A(x^2)^{(K-1)/2} & \text{if } K \text{ is odd.} \end{cases}$$

*Proof.* Assume first $K$ is even. Since $b^{\mathrm{PL}}(N, K) = 0$ when $N$ is odd, we get

$$B_K^{\mathrm{PL}}(x) = \sum_{s=1}^{\infty} b^{\mathrm{L}}\Big(s, \frac{K}{2}\Big)\,(x^2)^s = A(x^2)^{K/2}.$$

Assume $K$ is odd. If $K = 1$, the result is trivial because $B_{K=1}^{\mathrm{PL}}(x) = A(x)$. If $K \geq 3$, then

$$\begin{aligned} B_K^{\mathrm{LP}}(x) &= \sum_{m=1}^{\infty}\sum_{i=1}^{m-1} a(2i)\, b^{\mathrm{L}}\Big(m-i, \frac{K-1}{2}\Big)\, x^{2m} + \sum_{m=0}^{\infty}\sum_{i=1}^{m} a(2i-1)\, b^{\mathrm{L}}\Big(m+1-i, \frac{K-1}{2}\Big)\, x^{2m+1} \\ &= \sum_{i=1}^{\infty} a(2i)x^{2i} \sum_{m=i+1}^{\infty} b^{\mathrm{L}}\Big(m-i, \frac{K-1}{2}\Big)\, x^{2(m-i)} \\ &\qquad + \sum_{i=1}^{\infty} a(2i-1)x^{2i-1} \sum_{m=i}^{\infty} b^{\mathrm{L}}\Big(m-i+1, \frac{K-1}{2}\Big)\, x^{2(m-i)+2}. \end{aligned}$$

If we let $A_{\mathrm{E}}(x) = \sum_{i=1}^{\infty} a(2i)x^{2i}$ and $A_{\mathrm{O}}(x) = \sum_{i=1}^{\infty} a(2i-1)x^{2i-1}$, we then get

$$\begin{aligned} B_K^{\mathrm{LP}}(x) &= A_{\mathrm{E}}(x) \sum_{\ell=1}^{\infty} b^{\mathrm{L}}\Big(\ell, \frac{K-1}{2}\Big)\, x^{2\ell} + A_{\mathrm{O}}(x) \sum_{\ell=0}^{\infty} b^{\mathrm{L}}\Big(\ell+1, \frac{K-1}{2}\Big)\, x^{2(\ell+1)} \\ &= A_{\mathrm{E}}(x)A(x^2)^{(K-1)/2} + A_{\mathrm{O}}(x)A(x^2)^{(K-1)/2} = A(x)A(x^2)^{(K-1)/2}. \end{aligned}$$

□

**Theorem 2.2.** *The g.f. of the sequence $(b^{\mathrm{PL}}(N) : N \in \mathbb{Z}_{>0})$ is given by*

$$B^{\mathrm{LP}}(x) = \sum_{N=1}^{\infty} b^{\mathrm{PL}}(N)\, x^N = \frac{A(x) + A(x^2)}{1 - A(x^2)}.$$

*Proof.* By Proposition 1.1,

$$\sum_{N=1}^{\infty} b^{\mathrm{PL}}(N)\, x^N = \sum_{K=1}^{\infty}\left(\sum_{N=1}^{\infty} b^{\mathrm{PL}}(N, K)\, x^N\right)$$
$$= \sum_{s=1}^{\infty} A(x^2)^{2s/2} + \sum_{s=0}^{\infty} A(x)A(x^2)^{((2s+1)-1)/2} = \frac{A(x)+A(x^2)}{1-A(x^2)}. \qquad \square$$

**Example 2.3.** Consider again the $m$-colored compositions from Example 1.2 with input sequence $a = (m : m \in \mathbb{Z}_{>0})$ and input g.f. $A(x) = x/(1-x)^2$. Using Theorems 2.1 and 2.2, we can easily prove that

$$B_K^{\mathrm{PL}}(x) = \frac{x^K}{(1-x^2)^K} \qquad \text{for } K \text{ even},$$
$$B_K^{\mathrm{PL}}(x) = \frac{x^K(1+x)^2}{(1-x^2)^{K+1}} \qquad \text{for } K \text{ odd},$$
$$B^{\mathrm{PL}}(x) = \frac{x(x^2+3x+1)}{x^4-3x^2+1} = \frac{x(x^2+3x+1)}{(x^2+x-1)(x^2-x-1)}.$$

It follows that the number of $m$-colored (linear) palindromic compositions of $N$ with $K$ parts is given by

$$b^{\mathrm{PL}}(N, K) = \begin{cases} \frac{1+(-1)^N}{2}\binom{\lfloor \frac{N}{2}\rfloor + \frac{K}{2} - 1}{K-1} & \text{if } K \text{ is even}, \\ \binom{\lfloor \frac{N}{2}\rfloor + \frac{K-1}{2}}{K} + \binom{\lceil \frac{N}{2}\rceil + \frac{K-1}{2}}{K} & \text{if } K \text{ is odd}. \end{cases}$$

Here, $\lfloor a \rfloor$ and $\lceil a \rceil$ denote the floor and ceiling of $a \in \mathbb{R}$, respectively. In addition, the total number of $m$-colored (linear) palindromic compositions of $N$ is given by

$$b^{\mathrm{PL}}(N) = \begin{cases} 3F_N & \text{if } N \text{ is even}, \\ F_{N-1} + F_{N+1} & \text{if } N \text{ is odd}. \end{cases}$$

**Example 2.4.** Using a combinatorial argument and using g.f.'s, Mansour and Shattuck [2014] proved that, for $N \in \mathbb{Z}_{>0}$, the number of $F_m$-compositions of $N$ equals the *Pell number* $p_N$, which is defined by the recurrence

$$p_0 = 0, \qquad p_1 = 1, \qquad \text{and} \qquad p_N = 2p_{N-1} + p_{N-2} \quad \text{for } N \geq 2. \tag{2-1}$$

Of course, using g.f.'s, the proof of this claim is very easy: we have $a(m) = F_m$ for each $m \in \mathbb{Z}_{>0}$, and as a result, $A(x) = x/(1-x-x^2)$. It follows that the g.f. of the number $b^{\mathrm{L}}(N)$ of colored linear compositions of $N$ with respect to the input sequence $a = (F_m : m \in \mathbb{Z}_{>0})$ is

$$B^{\mathrm{L}}(x) = \frac{A(x)}{1-A(x)} = \frac{x}{1-2x-x^2},$$

which is the g.f. of the Pell numbers defined by (2-1).

By Theorem 2.2, the g.f. of the number $b^{\mathrm{LP}}(N)$ of colored linear palindromic compositions of $N$ with respect to the input sequence $a = (F_m : m \in \mathbb{Z}_{>0})$ is

$$B^{\mathrm{LP}}(x) = \frac{A(x)+A(x^2)}{1-A(x^2)} = \frac{x(1+x-2x^2-x^3-x^4)}{(1-x-x^2)(1-2x^2-x^4)} = \frac{2(x+1)}{1-x-x^2} - \frac{2+3x+x^3}{1-2x^2-x^4}.$$

It follows immediately that

$$b^{\mathrm{LP}}(N) = \begin{cases} 2F_{N+2} - 2p_{N/2+1} & \text{if } N \text{ is even,} \\ 2F_{N+2} - 3p_{(N+1)/2} - p_{(N-1)/2} & \text{if } N \text{ is odd} \end{cases} = 2F_{N+2} - p_{\lfloor N/2+1 \rfloor} - p_{\lceil N/2+1 \rceil}.$$

For example, the $b^{\mathrm{LP}}(5) = 9$ linear palindromic $F_m$-compositions of $N = 5$ are

$$(5_1),\ (5_2),\ (5_3),\ (5_4),\ (5_5),\ (1_1, 3_1, 1_1),\ (1_1, 3_2, 1_1),\ (2_1, 1_1, 2_1),\ (1_1, 1_1, 1_1, 1_1, 1_1).$$

**Remark 2.5.** Hadjicostas and Zhang [2017] considered also *type-II palindromic compositions* of $N$ with $K$, denoted by $P^{\mathrm{L}_2}(N, K)$, which are compositions $(\lambda_1, \ldots, \lambda_K)$ of $N$ of length $K$ that satisfy $\lambda_i = \lambda_{K+2-i}$ for $i = 2, \ldots, K$; that is, $(\lambda_1, \lambda_2, \ldots, \lambda_K) = (\lambda_1, \lambda_K, \ldots, \lambda_2)$. (For $K = 1$, it is assumed that $(\lambda_1) = (N)$ is a linear palindromic composition of both types.) Again, we shall not use the notation $P^{\mathrm{L}_2}(N, K)$ in this paper (since the superscript L in this paper denotes a general linear composition, not necessarily palindromic).

## 3. Colored circular palindromic compositions

*Circular palindromic compositions* or *circular symmetrical compositions* were originally studied in [Sommerville 1909]. Hadjicostas and Zhang [2017] defined them as equivalence classes (with respect to cyclic shifts) on the set of linear compositions of $N$ with $K$ parts that contain at least one palindromic composition of type I or type II (see Remark 2.5 in this paper). Williamson [1972] called them *bilaterally symmetric cyclic compositions*, while Bower [2010] called them *circular palindromes*.

In this paper, we use the superscript CP for colored circular palindromes (as opposed to PL used for colored linear palindromes and L for general colored linear compositions). In order to find the g.f.s of the sequences $(b^{\mathrm{CP}}(N, K) : N \in \mathbb{Z}_{>0})$ and $(b^{\mathrm{CP}}(N) : N \in \mathbb{Z}_{>0})$, we need first to express these quantities in terms of $b^{\mathrm{L}}(N, K)$ or $b^{\mathrm{LP}}(N, K)$ (for which we know the g.f.'s from previous sections).

Using input sequence $a = (a(m) : m \in \mathbb{Z})$, for $N, K \in \mathbb{Z}_{>0}$, it is clear that

$$b^{\mathrm{CP}}(N, K) = b^{\mathrm{PL}}(N, K) \quad \text{when } K \text{ is odd.} \tag{3-1}$$

When $K$ is even and $N$ is odd, we have (e.g., see [Bower 2010])

$$b^{\mathrm{CP}}(N, K) = \begin{cases} \sum_{\substack{i \text{ odd}, j \text{ even} \\ i+j=N}} a(i)\, a(j) & \text{if } K = 2, \\ \sum_{\substack{i \text{ odd}, j \text{ even} \\ i+j<N}} a(i)\, a(j)\, b^{\mathrm{L}}\Big(\frac{N-i-j}{2}, \frac{K}{2} - 1\Big) & \text{if } K \geq 4. \end{cases} \tag{3-2}$$

When $N$ and $K$ are both even, the situation is more complicated. Following [Bower 2010], we divide the possible output configurations (with parts of the composition on a circle) into two kinds:

Case 1: those configurations for which the axis of symmetry passes through no parts of the composition.

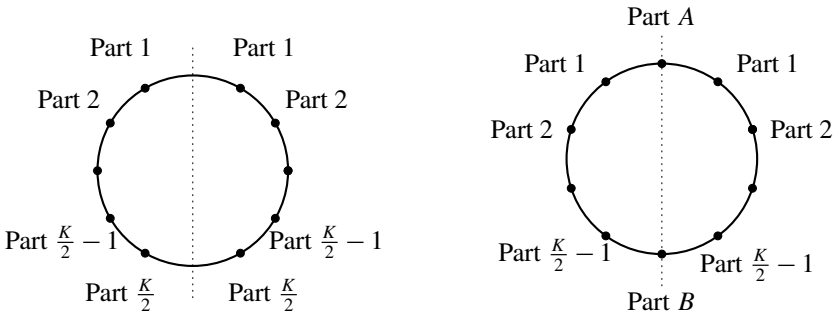



**Figure 1.** Case 1, left: no parts joined. Case 2, right: two parts joined.

Case 2: those configurations for which the axis of symmetry passes through two parts, which we label Part A and Part B.

See Figure 1. Because two circular configurations of parts are equivalent if one can be obtained from the other through cyclic rotation, it is possible for a configuration to belong to both categories.

We have the following formula for $b^{\mathrm{CP}}(N, K)$ when both $N$ and $K$ are even:

$$b^{\mathrm{CP}}(N, K) = \frac{I_N + J_N}{2} + P_N + S_N + M_N.$$

The quantities $I_N$, $J_N$, $P_N$, $S_N$, and $M_N$ are defined by [Bower 2010] (see his CPAL transform):

- $I_N = b^{\mathrm{L}}(N/2, K/2)$ (no parts are joined).
- $J_N = a(N/2)$ when $K = 2$ and

$$J_N = \sum_{0<i<N/2} a(i)\, b^{\mathrm{L}}\Big(\frac{N-2i}{2}, \frac{K}{2} - 1\Big)$$

  when $K \geq 4$ (the two parts joined, $A$ and $B$, are identical).

- $P_N = \begin{cases} \sum_{\substack{i,j \text{ even} \\ j>i,\, i+j=N}} a(i)\, a(j) & \text{when } K = 2, \\ \sum_{\substack{i,j \text{ even} \\ j>i,\, i+j<N}} a(i)\, a(j)\, b^{\mathrm{L}}\Big(\frac{N-i-j}{2}, \frac{K}{2} - 1\Big) & \text{when } K \geq 4 \end{cases}$

  (the two parts joined, $A$ and $B$, are even and have different values).

- $S_N = \begin{cases} \sum_{\substack{i,j \text{ odd} \\ j>i,\, i+j=N}} a(i)\, a(j) & \text{when } K = 2, \\ \sum_{\substack{i,j \text{ odd} \\ j>i,\, i+j<N}} a(i)\, a(j)\, b^{\mathrm{L}}\Big(\frac{N-i-j}{2}, \frac{K}{2} - 1\Big) & \text{when } K \geq 4 \end{cases}$

  (the two parts joined, $A$ and $B$, are odd and have different values).

- $M_N = \begin{cases} \frac{1}{2}\Big(a\Big(\frac{N}{2}\Big)^2 - a\Big(\frac{N}{2}\Big)\Big) & \text{when } K = 2, \\ \frac{1}{2}\sum_{0<i<N/2} (a(i)^2 - a(i))\, b^{\mathrm{L}}\Big(\frac{N-2i}{2}, \frac{K}{2} - 1\Big) & \text{when } K \geq 4 \end{cases}$

  (the two parts joined, $A$ and $B$, have the same value but different colors).

**Theorem 3.1.** *For fixed $K \in \mathbb{Z}_{>0}$, the g.f. of the sequence $(b^{\mathrm{CP}}(N, K) : N \in \mathbb{Z}_{>0})$ is given by*

$$B_K^{\mathrm{CP}}(x) = \sum_{N=1}^{\infty} b^{\mathrm{CP}}(N, K)\, x^N = \begin{cases} \frac{1}{2} A(x^2)^{(K/2)-1}(A(x)^2 + A(x^2)) & \text{if } K \text{ is even,} \\ A(x) A(x^2)^{(K-1)/2} & \text{if } K \text{ is odd.} \end{cases}$$

*Proof.* In view of (3-1) and Theorem 2.1, we only need to prove the theorem when $K$ is even. Let $B_K^{\mathrm{O}}(x)$ and $B_K^{\mathrm{E}}(x)$ be the contribution to the g.f. $B_K^{\mathrm{CP}}(x) = \sum_{N=1}^{\infty} b^{\mathrm{CP}}(N, K)\, x^N$ from the terms $b^{\mathrm{CP}}(N, K)$ with odd and even indexes $N$, respectively. We claim that

$$B_K^{\mathrm{O}}(x) = A_{\mathrm{O}}(x) A_{\mathrm{E}}(x) A(x^2)^{K/2-1}, \tag{3-3}$$

$$B_K^{\mathrm{E}}(x) = \tfrac{1}{2} A(x^2)^{K/2-1}[A(x^2) + A_{\mathrm{E}}(x)^2 + A_{\mathrm{O}}(x)^2], \tag{3-4}$$

where $A_{\mathrm{O}}(x) = \sum_{m=0}^{\infty} a(2m+1)\, x^{2m+1}$ and $A_{\mathrm{E}}(x) = \sum_{m=1}^{\infty} a(2m)\, x^{2m}$. From (3-3) and (3-4), we get

$$B_K^{\mathrm{CP}}(x) = B_K^{\mathrm{O}}(x) + B_K^{\mathrm{E}}(x) = \tfrac{1}{2} A(x^2)^{K/2-1}[A(x^2) + A(x)^2],$$

and this would complete the proof of theorem.

Proof of equation (3-3): We use (3-2). For $K = 2$, we have

$$B_{K=2}^{\mathrm{O}}(x) = \sum_{t=1}^{\infty} \sum_{\substack{s \geq 0, r \geq 1 \\ s+r=t}} a(2s+1)\, a(2r)\, x^{2t+1} = \sum_{s=0}^{\infty} \sum_{r=1}^{\infty} a(2s+1)\, a(2r)\, x^{2r} x^{2s+1} = A_{\mathrm{O}}(x) A_{\mathrm{E}}(x) A(x^2)^{2/2-1}.$$

For $K$ even $\geq 4$, we have

$$\begin{aligned} B_K^{\mathrm{O}}(x) &= \sum_{t=2}^{\infty} \sum_{s=0}^{t-2} \sum_{r=1}^{t-1-s} a(2s+1)\, a(2r)\, b^{\mathrm{L}}\Big(t-r-s, \frac{K}{2}-1\Big)\, x^{2t+1} \\ &= \sum_{s=0}^{\infty} \sum_{r=1}^{\infty} a(2s+1)\, a(2r)\, x^{2r} x^{2s+1} \sum_{t=r+s+1}^{\infty} b^{\mathrm{L}}\Big(t-r-s, \frac{K}{2}-1\Big)\, x^{2(t-r-s)} \\ &= A_{\mathrm{O}}(x) A_{\mathrm{E}}(x) A(x^2)^{K/2-1}, \end{aligned}$$

which proves (3-3).

Proof of equation (3-4): We calculate the contributions of the terms $I_N$, $J_N$, $P_N$, $S_N$, and $M_N$ to the generating function $B_K^{\mathrm{E}}(x)$. For $T \in \{I, J, P, S, M\}$, denote the corresponding contribution to the g.f. $B_K^{\mathrm{E}}(x)$ by $B_K^T(x)$. We claim that

$$\begin{aligned} B_K^I(x) &= \sum_{m=1}^{\infty} I_{2m} x^{2m} = A(x^2)^{K/2}, \\ B_K^J(x) &= \sum_{m=1}^{\infty} J_{2m} x^{2m} = \Big(\sum_{i=1}^{\infty} a(i) x^{2i}\Big) A(x^2)^{K/2-1}, \\ B_K^P(x) &= \sum_{m=1}^{\infty} P_{2m} x^{2m} = \Big(\sum_{i=1}^{\infty} \sum_{j=i+1}^{\infty} a(2i)\, a(2j)\, x^{2(i+j)}\Big) A(x^2)^{K/2-1}, \\ B_K^S(x) &= \sum_{m=1}^{\infty} S_{2m} x^{2m} = \Big(\sum_{i=0}^{\infty} \sum_{j=i+1}^{\infty} a(2i+1)\, a(2j+1)\, x^{2(i+j+1)}\Big) A(x^2)^{K/2-1}, \\ B_K^M(x) &= \sum_{m=1}^{\infty} M_{2m} x^{2m} = \frac{1}{2}\Big(\sum_{i=1}^{\infty} (a(i)^2 - a(i))\, x^{2i}\Big) A(x^2)^{K/2-1}. \end{aligned}$$

Below we prove the formulae for $B_K^J(x)$ and $B_K^S(x)$. The proofs for the rest are similar, and hence we omit them.

For $B_K^J(x)$, when $K=2$, we have

$$B_{K=2}^J(x)=\sum_{m=0}^{\infty}J_{2m}x^{2m}=\sum_{m=0}^{\infty}a\Big(\frac{2m}{2}\Big)x^{2m}=\bigg(\sum_{i=1}^{\infty}a(i)x^{2i}\bigg)A(x^2)^{2/2-1}.$$

For $K$ even $\geq 4$, we have

$$\begin{aligned}B_K^J(x)&=\sum_{m=1}^{\infty}J_{2m}x^{2m}=\sum_{m=1}^{\infty}\bigg(\sum_{0<i<m}a(i)\,b^{\mathrm{L}}\Big(\frac{2m-2i}{2},\,\frac{K}{2}-1\Big)\bigg)x^{2m}\\&=\sum_{i=1}^{\infty}a(i)x^{2i}\sum_{m=i+1}^{\infty}b^{\mathrm{L}}\Big(m-i,\,\frac{K}{2}-1\Big)\,x^{2(m-i)}\\&=\sum_{i=1}^{\infty}a(i)x^{2i}\sum_{\ell=1}^{\infty}b^{\mathrm{L}}\Big(\ell,\,\frac{K}{2}-1\Big)x^{2\ell}=\bigg(\sum_{i=1}^{\infty}a(i)x^{2i}\bigg)A(x^2)^{K/2-1}.\end{aligned}$$

For $B_K^S(x)$, when $K=2$, we have

$$\begin{aligned}B_{K=2}^S(x)&=\sum_{m=1}^{\infty}S_{2m}x^{2m}=\sum_{m=2}^{\infty}\sum_{\substack{s>t\geq 0\\ s+t=m-1}}a(2t+1)a(2s+1)x^{2m}\\&=\bigg(\sum_{t=0}^{\infty}\sum_{s=t+1}^{\infty}a(2s+1)\,a(2t+1)\,x^{2(s+t+1)}\bigg)A(x^2)^{2/2-1}.\end{aligned}$$

For $K$ even $\geq 4$, we have

$$\begin{aligned}B_K^S(x)&=\sum_{m=1}^{\infty}S_{2m}x^{2m}=\sum_{m=1}^{\infty}\bigg(\sum_{\substack{i,\,j\text{ odd}\\ j>i,\,i+j<2m}}a(i)\,a(j)\,b^{\mathrm{L}}\Big(\frac{2m-i-j}{2},\,\frac{K}{2}-1\Big)\bigg)x^{2m}\\&=\sum_{t=0}^{\infty}\sum_{\ell=t+1}^{\infty}\sum_{m=t+\ell+2}^{\infty}a(2t+1)\,a(2\ell+1)\,b^{\mathrm{L}}\Big(m-t-\ell-1,\,\frac{K}{2}-1\Big)x^{2m}.\end{aligned}$$

Therefore,

$$\begin{aligned}B_K^S(x)&=\sum_{t=0}^{\infty}\sum_{\ell=t+1}^{\infty}a(2t+1)\,a(2\ell+1)\,x^{2(t+\ell+1)}\\&\qquad\times\sum_{m=t+\ell+2}^{\infty}a(2t+1)\,a(2\ell+1)\,b^{\mathrm{L}}\Big(m-t-\ell-1,\,\frac{K}{2}-1\Big)\,x^{2(m-t-\ell-1)}\\&=\sum_{t=0}^{\infty}\sum_{\ell=t+1}^{\infty}a(2t+1)\,a(2\ell+1)\,x^{2(t+\ell+1)}\sum_{s=1}^{\infty}b^{\mathrm{L}}\Big(s,\,\frac{K}{2}-1\Big)\,x^{2s}\\&=\bigg(\sum_{i=0}^{\infty}\sum_{j=i+1}^{\infty}a(2i+1)\,a(2j+1)\,x^{2(i+j+1)}\bigg)A(x^2)^{K/2-1}.\end{aligned}$$

In a similar way, we can prove the formulae for $B_K^I(x)$, $B_K^P(x)$, and $B_K^M(x)$. We then have

$$\begin{aligned} B_K^{\mathrm{E}}(x) &= \tfrac{1}{2}(B_K^I(x) + B_K^J(x)) + B_K^P(x) + B_K^S(x) + B_K^M(x) \\ &= A(x^2)^{K/2-1} T_K(x), \end{aligned}$$

where

$$\begin{aligned} T_K(x) &= \frac{1}{2}\Big(A(x^2) + \sum_{i=1}^{\infty} a(i)x^{2i}\Big) + \sum_{i=1}^{\infty}\sum_{j=i+1}^{\infty} a(2i)\,a(2j)\,x^{2(i+j)} \\ &\qquad + \sum_{i=0}^{\infty}\sum_{j=i+1}^{\infty} a(2i+1)\,a(2j+1)\,x^{2(i+j+1)} + \frac{1}{2}\sum_{i=1}^{\infty}(a(i)^2 - a(i))\,x^{2i} \\ &= \frac{1}{2}\Big(A(x^2) + \sum_{i=1}^{\infty} a(i)x^{2i} + A_{\mathrm{E}}(x)^2 - \sum_{m=1}^{\infty} a(2m)^2 x^{4m} + A_{\mathrm{O}}(x)^2 \\ &\qquad - \sum_{m=0}^{\infty} a(2m+1)^2 x^{2(2m+1)} + \sum_{i=1}^{\infty}(a(i)^2 - a(i))\,x^{2i}\Big) \\ &= \frac{1}{2}(A(x^2) + A_{\mathrm{E}}(x)^2 + A_{\mathrm{O}}(x)^2). \end{aligned}$$

This finishes the proof of (3-4) and the proof of the theorem. □

**Theorem 3.2.** *The g.f. of the sequence* $(b^{\mathrm{CP}}(N) : N \in \mathbb{Z}_{>0})$ *is given by*

$$B^{\mathrm{CP}}(x) = \sum_{N=1}^{\infty} b^{\mathrm{CP}}(N)\,x^N = \frac{(1+A(x))^2}{2(1-A(x^2))} - \frac{1}{2}.$$

*Proof.* By Proposition 1.1,

$$\begin{aligned} \sum_{N=1}^{\infty} b^{\mathrm{CP}}(N)\,x^N &= \sum_{K=1}^{\infty}\Big(\sum_{N=1}^{\infty} b^{\mathrm{CP}}(N,K)\,x^N\Big) \\ &= \frac{1}{2}\sum_{s=1}^{\infty} A(x^2)^{(2s/2)-1}(A(x)^2 + A(x^2)) + \sum_{s=0}^{\infty} A(x)A(x^2)^{(2s+1-1)/2} \\ &= \frac{A(x)^2 + A(x^2) + 2A(x)}{2(1-A(x^2))} = \frac{(1+A(x))^2}{2(1-A(x^2))} - \frac{1}{2}. \end{aligned}$$

□

A special case of Theorems 3.1 and 3.2 has to do with circular palindromic compositions of a positive integer with parts belonging to a subset $E$ of $\mathbb{Z}_{>0}$. (Several similar results for various kinds of linear compositions were surveyed in [Heubach and Mansour 2004].) Hadjicostas and Zhang [2017, Theorem 2.6 and Corollary 2.9] proved the following result using different methods. We give a new proof of this result using Theorems 3.1 and 3.2 above.

**Theorem 3.3.** *Let* $E \subseteq \mathbb{Z}_{>0}$. *For each pair of positive integers* $N$ *and* $K$, *let* $P_E^R(N,K)$ *be the number of circular palindromic compositions of* $N$ *with length* $K$ *whose parts belong to* $E$. *Let also* $P_E^R(N)$ *be the total number of circular palindromic compositions of* $N$ *with parts in* $E$. *Then*:

(a) *For $K \in \mathbb{Z}_{>0}$,*

$$\sum_{N\geq 1} P_E^R(N,K)\,x^N = \begin{cases} \Big(\sum_{m\in E} x^m\Big)\Big(\sum_{m\in E} x^{2m}\Big)^{(K-1)/2} & \text{if } K \text{ is odd,} \\ \frac{1}{2}\Big(\Big(\sum_{m\in E} x^m\Big)^2 + \Big(\sum_{m\in E} x^{2m}\Big)\Big)\Big(\sum_{m\in E} x^{2m}\Big)^{(K)/2-1} & \text{if } K \text{ is even.} \end{cases}$$

(b) *We have*

$$\sum_{N\geq 1} P_A^R(N)\,x^N = \frac{\big(1+\sum_{m\in E} x^m\big)^2}{2\big(1-\sum_{m\in E} x^{2m}\big)} - \frac{1}{2}.$$

*Proof.* Define the input sequence $a = (a(m) : m \geq 1)$ by

$$a(m) = \begin{cases} 1 & \text{if } m \in E, \\ 0 & \text{if } m \notin E. \end{cases}$$

Then the g.f. of sequence $a$ is $A(x) = \sum_{x\in E} x^m$. We have

$$P_E^R(N;K) = b^{\mathrm{CP}}(N,K) \quad \text{and} \quad P_E^R(N) = b^{\mathrm{CP}}(N).$$

The theorem then follows from Theorems 3.1 and 3.2 above. □

**Example 3.4.** Consider again the $m$-colored compositions from Example 1.2 with input sequence $a = (m : m \in \mathbb{Z}_{>0})$ and input g.f. $A(x) = x/(1-x)^2$. Using Theorems 3.1 and 3.2, we can easily prove that

$$B_K^{\mathrm{CP}}(x) = \frac{x^K(1+x^2)(1+x)^2}{(1-x^2)^{K+2}} \qquad \text{for } K \text{ even,}$$

$$B_K^{\mathrm{CP}}(x) = \frac{x^K(1+x)^2}{(1-x^2)^{K+1}} \qquad \text{for } K \text{ odd,}$$

$$B^{\mathrm{CP}}(x) = \frac{x(x^4+x^3-2x^2+x+1)}{(x^4-3x^2+1)(1-x)^2} = \frac{x(x^4+x^3-2x^2+x+1)}{(x^2+x-1)(x^2-x-1)(1-x)^2}.$$

Thus, the number of $m$-colored circular palindromic compositions of $N$ with $K$ parts is given by

$$b^{\mathrm{CP}}(N,K) = \begin{cases} \binom{\lfloor \frac{N-1}{2} \rfloor + \frac{K}{2}}{K+1} + \binom{\lceil \frac{N-1}{2} \rceil + \frac{K}{2}}{K+1} + \binom{\lfloor \frac{N+1}{2} \rfloor + \frac{K}{2}}{K+1} + \binom{\lceil \frac{N+1}{2} \rceil + \frac{K}{2}}{K+1} & \text{if } K \text{ is even,} \\ \binom{\lfloor \frac{N}{2} \rfloor + \frac{K-1}{2}}{K} + \binom{\lceil \frac{N}{2} \rceil + \frac{K-1}{2}}{K} & \text{if } K \text{ is odd.} \end{cases}$$

In addition, the total number of $m$-colored circular palindromic compositions of $N$ is given by

$$b^{\mathrm{CP}}(N) = F_{N+4} + (-1)^N F_{N-4} - 2N \quad \text{for } N \geq 4$$

with $b^{\mathrm{CP}}(1) = 1$, $b^{\mathrm{CP}}(2) = 3$ and $b^{\mathrm{CP}}(3) = 6$.

**Example 3.5.** Consider again Example 2.4 with input sequence $a = (F_m : m \in \mathbb{Z}_{>0})$ and input g.f. $A(x) = x/(1-x-x^2)$, which extends an example from [Mansour and Shattuck 2014]. Using Theorem 3.2,

we can prove that the g.f. of the number $b^{\mathrm{CP}}(N)$ of colored circular palindromic $F_m$-compositions of $N$ is given by

$$B^{\mathrm{CP}}(x) = \frac{x(1-3x^2+x^4+x^5+x^6)}{(1-x-x^2)^2(1-2x^2-x^4)} = -\frac{5(x+2)}{1-x-x^2} + \frac{1}{(1-x-x^2)^2} + \frac{9+14x+4x^2+6x^3}{1-2x^2-x^4}.$$

It follows that

$$b^{\mathrm{CP}}(N) = -5(F_N+2F_{N+1}) + \tfrac{1}{5}((N+3)F_{N+1}+(N+1)F_{N+3}) + g(N),$$

where

$$g(N) = \begin{cases} 4p_{N/2}+9p_{N/2+1} & \text{if } N \text{ is even,} \\ 6p_{(N-1)/2}+14p_{(N+1)/2} & \text{if } N \text{ is odd.} \end{cases}$$

Here $p_n$ denotes the $n$-th Pell number defined by (2-1). For example, the $b^{\mathrm{CP}}(5) = 15$ circular palindromic $F_m$-compositions of $N = 5$ are

$$(5_1),\ (5_2),\ (5_3),\ (5_4),\ (5_5),\ (1_1,4_1),\ (1_1,4_2),\ (1_1,4_3),\ (2_1,3_1),\ (2_1,3_2),$$
$$(1_1,3_1,1_1),\ (1_1,3_2,1_1),\ (2_1,1_1,2_1),\ (1_1,1_1,1_1,2_1),\ (1_1,1_1,1_1,1_1,1_1),$$

where we have listed only one representative from each equivalence class.

## Acknowledgements

We would like to thank the two anonymous referees whose comments and suggestions helped us improve the presentation of the paper.

## References

[Abrate et al. 2014] M. Abrate, S. Barbero, U. Cerruti, and N. Murru, "Colored compositions, invert operator and elegant compositions with the "black tie"", *Discrete Math.* **335** (2014), 1–7. MR Zbl

[Agarwal 2000] A. K. Agarwal, "$n$-colour compositions", *Indian J. Pure Appl. Math.* **31**:11 (2000), 1421–1427. MR Zbl

[Agarwal 2003] A. K. Agarwal, "An analogue of Euler's identity and new combinatorial properties of $n$-colour compositions", *J. Comput. Appl. Math.* **160**:1-2 (2003), 9–15. MR Zbl

[Bernstein and Sloane 1995] M. Bernstein and N. J. A. Sloane, "Some canonical sequences of integers", *Linear Algebra Appl.* **226/228** (1995), 57–72. Correction in **320** (2000), 210. MR Zbl

[Bower 2010] C. G. Bower, "Further transformations of integer sequences", web article, 2010, available at https://oeis.org/transforms2.html.

[Ferrari and Zagaglia Salvi 2018] M. M. Ferrari and N. Zagaglia Salvi, "Cyclic compositions and cycles of the hypercube", *Aequationes Math.* **92**:4 (2018), 671–682. MR Zbl

[Flajolet and Sedgewick 2009] P. Flajolet and R. Sedgewick, *Analytic combinatorics*, Cambridge University Press, 2009. MR Zbl

[Flajolet and Soria 1991] P. Flajolet and M. Soria, "The cycle construction", *SIAM J. Discrete Math.* **4**:1 (1991), 58–60. MR Zbl

[Gibson et al. 2018] M. M. Gibson, D. Gray, and H. Wang, "Combinatorics of $n$-color cyclic compositions", *Discrete Math.* **341**:11 (2018), 3209–3226. MR Zbl

[Guo 2012] Y.-h. Guo, "Some $n$-color compositions", *J. Integer Seq.* **15**:1 (2012), art. id. 12.1.2. MR Zbl

[Hadjicostas 2016] P. Hadjicostas, "Cyclic compositions of a positive integer with parts avoiding an arithmetic sequence", *J. Integer Seq.* **19**:8 (2016), art. id. 16.8.2. MR Zbl

[Hadjicostas 2017] P. Hadjicostas, "Cyclic, dihedral and symmetrical Carlitz compositions of a positive integer", *J. Integer Seq.* **20**:8 (2017), art. id. 17.8.5. MR Zbl

[Hadjicostas and Zhang 2017] P. Hadjicostas and L. Zhang, "Sommerville's symmetrical cyclic compositions of a positive integer with parts avoiding multiples of an integer", *Fibonacci Quart.* **55**:1 (2017), 54–73. MR Zbl

[Heubach and Mansour 2004] S. Heubach and T. Mansour, "Compositions of $n$ with parts in a set", *Congr. Numer.* **168** (2004), 127–143. MR Zbl

[Heubach and Mansour 2010] S. Heubach and T. Mansour, *Combinatorics of compositions and words*, CRC Press, Boca Raton, FL, 2010. MR Zbl

[Knopfmacher and Robbins 2010] A. Knopfmacher and N. Robbins, "Some properties of cyclic compositions", *Fibonacci Quart.* **48**:3 (2010), 249–255. MR Zbl

[Knopfmacher and Robbins 2013] A. Knopfmacher and N. Robbins, "Some properties of dihedral compositions", *Util. Math.* **92** (2013), 207–220. MR Zbl

[MacMahon 1893] P. A. MacMahon, "Memoir on the theory of the compositions of numbers", *Philos. Trans. Roy. Soc. London Ser. A* **184** (1893), 835–901. JFM

[Mansour and Shattuck 2014] T. Mansour and M. Shattuck, "A statistic on $n$-color compositions and related sequences", *Proc. Indian Acad. Sci. Math. Sci.* **124**:2 (2014), 127–140. MR Zbl

[Sachdeva and Agarwal 2017] R. Sachdeva and A. K. Agarwal, "Combinatorics of certain restricted $n$-color composition functions", *Discrete Math.* **340**:3 (2017), 361–372. MR Zbl

[Sommerville 1909] D. M. Y. Sommerville, "On certain periodic properties of cyclic compositions of numbers", *Proc. London Math. Soc.* (2) **7** (1909), 263–313. MR Zbl

[Williamson 1972] S. G. Williamson, "The combinatorial analysis of patterns and the principle of inclusion-exclusion", *Discrete Math.* **1**:4 (1972), 357–388. MR Zbl

[Zagaglia Salvi 1999] N. Zagaglia Salvi, "Ordered partitions and colourings of cycles and necklaces", *Bull. Inst. Combin. Appl.* **27** (1999), 37–40. MR Zbl

Received 5 Mar 2020. Revised 23 Jun 2020.

PETROS HADJICOSTAS:

peterhadji1@gmail.com
Department of Mathematical Sciences, University of Nevada, Las Vegas, NV, United States

msp

Moscow Journal of Combinatorics and Number Theory
Vol. 9, No. 2, 2020
https://doi.org/10.2140/moscow.2020.9.187

# Square-full primitive roots in arithmetic progressions

Vichian Laohakosol, Teerapat Srichan and Pinthira Tangsupphathawat

An asymptotic estimate for the number of positive primitive roots which are square-full integers in arithmetic progressions is derived. The employed method combines two techniques and is based on the character-sum method involving two characters; one character is to take care of being a primitive root, based on a result of Shapiro, and the other character is to take care of being square-full, based on a result of Munsch.

## 1. Introduction

An integer $n > 1$ is called square-full if in its canonical prime representation each prime appears with exponent $\geq 2$. The integer 1 is square-full by convention. For a positive real number $x$, let $Q_2(x)$ denote the number of square-full integers that are $\leq x$. The oldest known work related to $Q_2(x)$ is due to Erdős and Szekeres [1934], who proved that

$$Q_2(x) = \frac{\zeta\left(\frac{3}{2}\right)}{\zeta(3)} x^{\frac{1}{2}} + O(x^{\frac{1}{3}}).$$

This was later refined by Bateman and Grosswald [1958], who replaced the error term by

$$\frac{\zeta\left(\frac{2}{3}\right)}{\zeta(2)} x^{\frac{1}{3}} + O(x^{\frac{1}{6}} \exp(-C(\log x)^{\frac{3}{5}} (\log\log x)^{-\frac{1}{5}}))$$

for some absolute constant $C > 0$. There have been many works on the improvement of the error term; see, e.g., [Cai 1997; Cao 1994; 1997; Liu 1994; Suryanarayana and Sitaramachandra Rao 1973; Wu 1998; 2001]. Regarding square-full integers in an arithmetic progression, Liu and Zhang [2013] used Perron's formula and properties of the Dirichlet $L$-functions to study the character sums over square-full numbers and gave an asymptotic formula for $Q_2(x : \ell, q)$, the number of square-full integers which are congruent to $\ell$ modulo an integer $q$ and not exceeding $x$. One year later, Munsch [2014] applied the Pólya–Vinogradov inequality to bound the character sums over square-full integers and improved the results obtained by Liu and Zhang by showing that for all $\varepsilon > 0$ we have

$$Q_2(x; \ell, q) = \frac{\zeta\left(\frac{3}{2}\right)}{\zeta(3)} \frac{A_{\ell,q}}{q} x^{\frac{1}{2}} + \frac{\zeta\left(\frac{2}{3}\right)}{\zeta(2)} \frac{B_{\ell,q}}{q} x^{\frac{1}{3}} + O(x^{\frac{1}{4}} (\log x)^{\frac{1}{2}} q^{\frac{11}{32}+\varepsilon}), \tag{1-1}$$

where $A_{\ell,q}$, $B_{\ell,q}$ are constants depending on certain $L$-functions. Chan and Tsang [2013] used the Dirichlet hyperbola method and Burgess bound on character sums to study this problem. Later, Chan [2015] improved their results. Character sums over square-full integers are also prominent in [Liu and

*MSC2020:* 11A07, 11B25, 11N37, 11N69.
*Keywords:* arithmetic progressions, primitive roots, square-full numbers.

Zhang 2013; Munsch 2014]. The second author [Srichan 2013] used the exponent-pair method to study the appearance of square-full integers in an arithmetic progression and showed that the right-hand side of (1-1) can be strengthened to

$$\frac{L(\frac{3}{2},\chi_0)+\sum_{\chi_1}\bar{\chi}_1(\ell)L(\frac{3}{2},\chi_1)}{qL(3,\chi_0)}x^{\frac{1}{2}}+\frac{L(\frac{2}{3},\chi_0)+\sum_{\chi_2}\bar{\chi}_2(\ell)L(\frac{2}{3},\chi_2)}{qL(2,\chi_0)}x^{\frac{1}{3}}+O(q^{\frac{1}{2}+\varepsilon}x^{\frac{1}{6}}), \tag{1-2}$$

where $\chi_0$, $\chi_1$ and $\chi_2$ denote the principal, quadratic and cubic characters modulo $q$, respectively. Character sums over square-full integers play a significant role in the proof of (1-2).

For relatively prime integers $a, m$ with $m \geq 1$, if the smallest positive integer $f$ such that $a^f \equiv 1 \bmod m$ satisfies $f = \phi(m)$ (the Euler totient), then $a$ is called a primitive root mod $m$. Shapiro [1983, Sections 8.5–8.6] also used the character-sum method to obtain the following estimates related to the number of primitive roots modulo an odd prime $p$:

- The number of positive primitive roots mod $p$ that are $\leq x$ is

$$\frac{\phi(p-1)}{p-1}(x+O(2^{\omega(p-1)}p^{\frac{1}{2}}\log p)),$$

where $\omega(n)$ denotes the number of distinct prime factors of $n$.

- For integers $k > 0$, $\ell$ with $\gcd(p,k)=1$, the number of positive primitive roots mod $p$ that are $\leq x$ and $\equiv \ell \bmod k$ is

$$\frac{\phi(p-1)}{p-1}\left(\frac{x}{k}+O(2^{\omega(p-1)}p^{\frac{1}{2}}\log p)\right).$$

- The number of positive primitive roots mod $p$ that are $\leq x$ which are square-full is

$$\frac{\phi(p-1)}{p-1}(cx^{\frac{1}{2}}+O(x^{\frac{1}{3}}2^{\omega(p-1)})p^{\frac{1}{6}}(\log p)^{\frac{1}{3}}), \tag{1-3}$$

where $c$ is a constant.

Liu and Zhang [2005], using Perron's formula, improved the error term in (1-3) to $O(x^{1/4+\varepsilon}p^{9/44+\varepsilon})$. Munsch and Trudgian [2018] improved this result by showing that (1-3) can be replaced by

$$\frac{\phi(p-1)}{p-1}\left(\left(1+\frac{1}{p}+\frac{1}{p^2}\right)^{-1}\frac{C_p x^{\frac{1}{2}}}{\zeta(3)}+O(x^{\frac{1}{3}}(\log x)p^{\frac{1}{9}}(\log p)^{\frac{1}{6}}2^{\omega(p-1)})\right), \tag{1-4}$$

where $C_p \gg p^{-1/(8\sqrt{e})}$. Recently, Srichan [2020] used the exponent-pair method and the lemmas used in the proof of Theorem 2.1 in [Srichan 2013] to further refine the estimate (1-4) with the following result: for a given odd prime $p \leq x^{1/5}$, the number of square-full integers which are primitive roots mod $p$ and $\leq x$ is equal to

$$\frac{\phi(p-1)}{p}\left\{\left(\frac{L(\frac{3}{2},\chi_0)-L(\frac{3}{2},\chi_1)}{L(3,\chi_0)}\right)x^{\frac{1}{2}}+\left(\frac{L(\frac{2}{3},\chi_0)-L(\frac{2}{3},\chi_2^2)}{L(2,\chi_0)}\right)x^{\frac{1}{3}}\right\}$$
$$+O(x^{\frac{1}{6}}\phi(p-1)3^{\omega_{1,3}(p-1)}p^{\frac{1}{2}+\varepsilon}). \tag{1-5}$$

Here, $\chi_0$, $\chi_1 \neq \chi_0$, and $\chi_2 \neq \chi_0$ denote, respectively, the principal, quadratic and cubic characters mod $p$, and $\omega_{1,3}(n)$ denotes the number of distinct primes $\equiv 1 \pmod 3$ which are divisors of $n$.

In the present work, we derive an asymptotic estimate for the number of primitive roots mod $p$ which are square-full in an arithmetic progression. Shapiro [1983] was first to give an asymptotic formula for the number of primitive roots mod $p$ in an arithmetic progression by showing for given integers $k > 0$, $\ell$, and prime $p$ with $\gcd(p, k) = 1$, the number of positive primitive roots modulo $p$ that are congruent to $\ell$ mod $k$ and not exceeding $x$ is equal to

$$\frac{\phi(p-1)}{p-1}\left(\frac{x}{k} + O(2^{\omega(p-1)}p^{\frac{1}{2}}\log p)\right).$$

Throughout, let $\varepsilon$ be a fixed sufficiently small positive constant, let $\phi(n)$ be Euler's totient, let $\mu(n)$ be the Möbius function, and for a given odd prime $p$ let

$$T_2(n) = \begin{cases} 1 & \text{if } n \text{ is a square-full primitive root mod } p, \\ 0 & \text{otherwise} \end{cases} \tag{1-6}$$

be the characteristic function of the primitive roots modulo $p$ which are square-full integers. Our main result is:

**Theorem 1.1.** *Given an integer $q \geq 2$, an integer $0 < \ell < q$ with $\gcd(\ell, q) = 1$, and a given odd prime $p$ such that $p \nmid q$, we have*

$$\sum_{\substack{n \leq x \\ n \equiv \ell \bmod q}} T_2(n) = \frac{\phi(p-1)}{\phi(p)\phi(q)}\big(A_{p,q}x^{\frac{1}{2}} + B_{p,q}x^{\frac{1}{3}} + O((pq)^{\frac{11}{32}+\varepsilon}x^{\frac{1}{4}}(\log x)^{\frac{1}{2}}2^{\omega(p-1)}\phi(q))\big),$$

*where*

$$A_{p,q} = \frac{\zeta(\frac{3}{2})}{\zeta(3)}\prod_{p_1 \mid pq}\left(\frac{1-p_1^{-1}}{1+p_1^{-\frac{3}{2}}}\right) - \sum_{\lambda \in Y_1}\frac{L(\frac{3}{2},\lambda)}{\zeta(3)}\left(1+\frac{1}{p}+\frac{1}{p^2}\right)^{-1}H_q(\tfrac{1}{2},\lambda) \tag{1-7}$$

$$+ \sum_{\chi \in X_1}\bar{\chi}(\ell)\frac{L(\frac{3}{2},\chi)}{\zeta(3)}\prod_{p_1 \mid q}\left(1+\frac{1}{p_1}+\frac{1}{p_1^2}\right)^{-1}F_p(\tfrac{1}{2},\chi) \tag{1-8}$$

$$- \sum_{\substack{\chi \in X_1 \\ \lambda \in Y_1}}\bar{\chi}(\ell)\frac{L(\frac{3}{2},\chi\lambda)}{\zeta(3)}\prod_{p_1 \mid pq}(1-p_1^{-3})^{-1}\frac{\phi(pq)}{pq}, \tag{1-9}$$

$$B_{p,q} = \frac{\zeta(\frac{2}{3})}{\zeta(2)}\prod_{p_1 \mid pq}\left(\frac{1-p_1^{-\frac{2}{3}}}{1+p_1^{-1}}\right) - \frac{1}{2}\sum_{\lambda \in Y_2}\frac{L(\frac{2}{3},\lambda^2)}{\zeta(2)}\left(1+\frac{1}{p}\right)^{-1}H_q(\tfrac{1}{3},\lambda) \tag{1-10}$$

$$+ \sum_{\chi \in X_2}\bar{\chi}(\ell)\frac{L(\frac{2}{3},\chi^2)}{\zeta(2)}\prod_{p_1 \mid q}\left(1+\frac{1}{p_1}\right)^{-1}F_p(\tfrac{1}{3},\chi) \tag{1-11}$$

$$- \frac{1}{2}\sum_{\substack{\chi \in X_2 \\ \lambda \in Y_2}}\bar{\chi}(\ell)\frac{L(\frac{2}{3},\chi^2\lambda^2)}{\zeta(2)}\prod_{p_1 \mid pq}(1+p_1^{-1}), \tag{1-12}$$

*with*

$$H_q(s,\lambda) = \prod_{p_1 \mid q}\frac{1-\lambda^2(p_1)p_1^{-2s}}{1+\lambda^3(p_1)p_1^{-3s}}, \quad F_p(s,\chi) = \frac{1-\chi^2(p)p^{-2s}}{1+\chi^3(p)p^{-3s}},$$

*the products being over primes* $p_1$. *Here,* $X_1, X_2$ *denote the set of quadratic, respectively, cubic characters mod* $q$, *while* $Y_1, Y_2$ *denote the set of quadratic, respectively, cubic characters mod* $p$.

Our approach is based mostly on that of [Munsch 2014]. However, in contrast to the analysis in that paper, which deals with character sums involving one character, here we work with character sums involving two characters; one to take care of square-full integers (the approach in [Munsch 2014]) and the other to take care of primitive roots (the approach in [Shapiro 1983]). This leads inevitably to handling many more subcases. The subcases with contributions towards the main terms arise from one of the characters being principal and are presented with detailed proofs, while proofs for those subcases with contributions only towards the error terms are given more tersely.

## 2. Lemmas

We collect in this section, several auxiliary results used in the proof of the theorem.

**Lemma 2.1** [Shapiro 1983, Lemma 8.5.1]. *For a given odd prime* $p$, *the characteristic function indicating if* $n$ *is a primitive root mod* $p$ *satisfies*

$$\frac{\phi(p-1)}{p-1}\sum_{d\mid p-1}\frac{\mu(d)}{\phi(d)}\sum_{\lambda\in\Gamma_d}\lambda(n)=\begin{cases}1 & \text{if } n \text{ is a primitive root mod } p,\\ 0 & \text{otherwise,}\end{cases}$$

*where* $\Gamma_d$ *denotes the set of characters of the character group mod* $p$ *that are of order* $d$.

The next lemma gives special cases of the well-known as Pólya–Vinogradov inequality, taken from [Iwaniec and Kowalski 2004, Theorem 12.5, p. 324].

**Lemma 2.2.** *For a real* $x \geq 1$, *an integer* $q \geq 2$, *and a Dirichlet character* $\chi$, *let*

$$S_\chi(x)=\sum_{1\leq n\leq x}\chi(n).$$

*For any nonprincipal character* $\chi$ *mod* $q$, *we have*:

(I) [Iwaniec and Kowalski 2004, Theorem 12.5, p. 324]

$$|S_\chi(x)|\leq 6\sqrt{q}\log q.$$

(II) [Burgess 1962; Iwaniec and Kowalski 2004, Theorem 12.6, p. 326]

$$|S_\chi(x)|\ll x^{\frac{1}{2}}q^{\frac{3}{16}+\varepsilon}. \tag{2-1}$$

(III) [Iwaniec and Kowalski 2004, equation (12.58)] *If* $q$ *is prime, then*

$$|S_\chi(x)|\ll x^{\frac{1}{2}}q^{\frac{3}{16}}(\log q)^{\frac{1}{2}}. \tag{2-2}$$

For two characters $\chi$ mod $q$ and $\lambda$ mod $p$, we define our main sum of the product of two characters over square-free integers:

$$V(x;\chi,\lambda):=\sum_{n\leq x}\mu(n)^2\chi(n)\lambda(n). \tag{2-3}$$

**Lemma 2.3.** *Let $\chi$ and $\lambda$ be nontrivial characters mod $q$ and mod $p$, respectively. We have*

$$|V(x;\chi,\lambda)| \ll \begin{cases} x^{\frac{1}{2}}(pq)^{\frac{1}{4}}(\log pq)^{\frac{1}{2}}, \\ x^{\frac{1}{2}}(\log x)(pq)^{\frac{3}{16}+\varepsilon}. \end{cases}$$

*Proof.* Since $\mu^2(n) = \sum_{d^2\mid n}\mu(d)$, we get

$$\begin{aligned} V(x;\chi,\lambda) &= \sum_{md^2\le x}\mu(d)\chi(m)\chi(d^2)\lambda(m)\lambda(d^2) \\ &= \sum_{d\le x^{1/2}}\mu(d)\chi^2(d)\lambda^2(d)\sum_{m\le x/d^2}\chi(m)\lambda(m). \end{aligned} \tag{2-4}$$

To obtain the first bound, we introduce $H \ge 1$ and split the sum into two parts

$$V(x;\chi,\lambda) = \sum_{d\le H}\mu(d)\chi^2(d)\lambda^2(d)\sum_{m\le x/d^2}\chi(m)\lambda(m) + \sum_{H<d\le x^{1/2}}\mu(d)\chi^2(d)\lambda^2(d)\sum_{m\le x/d^2}\chi(m)\lambda(m).$$

Using Lemma 2.2, since $\chi\cdot\lambda$ is a character mod $pq$, we have

$$|V(x;\chi,\lambda)| \ll \sum_{d\le H}\sqrt{pq}\,\log pq + \sum_{H<d\le x^{1/2}} x/d^2 \ll H\sqrt{pq}\log pq + \frac{x}{H}.$$

Choosing

$$H = \lfloor x^{\frac{1}{2}}(pq)^{-\frac{1}{4}}(\log pq)^{-\frac{1}{2}}\rfloor,$$

we obtain

$$|V(x;\chi,\lambda)| \ll x^{\frac{1}{2}}(pq)^{\frac{1}{4}}(\log pq)^{\frac{1}{2}},$$

which is the first bound. To obtain the second bound, we apply (2-1) in Lemma 2.2 to (2-4) to get

$$|V(x;\chi,\lambda)| \ll \sum_{d\le x^{1/2}}\left(\frac{x}{d^2}\right)^{\frac{1}{2}}(pq)^{\frac{3}{16}+\varepsilon} \ll x^{\frac{1}{2}}(\log x)(pq)^{\frac{3}{16}+\varepsilon}. \qquad \square$$

One can also obtain Lemma 2.3 by using Lemma 2.3 in [Munsch 2014] and the fact that if $\chi_1$ is a character mod $q_1$ and $\chi_2$ is a character mod $q_2$, their product $\chi_1\chi_2$ is a character mod $\operatorname{lcm}(q_1,q_2)$.

By proofs similar to those of Lemmas 2.5 and 2.7 in [Munsch 2014], we obtain:

**Lemma 2.4.** *Let $q$ be an integer $\ge 2$, let $p$ be an odd prime with $p\nmid q$, let $\chi,\lambda$ be nonprincipal characters mod $q$ and $p$, respectively, and let $H$ be a positive integer. Then*

$$\sum_{\substack{n\le x\\ \gcd(n,pq)=1}} 1 = \frac{\phi(pq)}{pq}x + O(\tau(pq)), \tag{2-5}$$

$$\sum_{\substack{n\le x\\ \gcd(n,pq)=1}} \mu^2(n) = \frac{x}{\zeta(2)}\prod_{p_1\mid pq}\left(1+\frac{1}{p_1}\right)^{-1} + O(x^{\frac{1}{2}}\tau(pq)), \tag{2-6}$$

$$\sum_{n\le H}\frac{\mu^2(n)\chi(n)\lambda(n)}{n^{\frac{3}{2}}} = \frac{L\left(\frac{3}{2},\chi\lambda\right)}{L(3,\chi^2\lambda^2)} + O(H^{-1}(\log H)(pq)^{\frac{3}{16}+\varepsilon}), \tag{2-7}$$

$$\sum_{n\le H}\frac{\chi(n)\lambda(n)}{n^{\frac{2}{3}}}=L\big(\tfrac{2}{3},\chi\lambda\big)+O(H^{-\frac{2}{3}}(\log pq)(pq)^{\frac{1}{2}}),\tag{2-8}$$

*where $\tau(n)$ denotes the number of positive divisors of $n$, and the product runs over primes $p_1\mid pq$.*

For any character $\chi \bmod q$, we define

$$Q_2(x;\chi):=\sum_{n\le x}\alpha(n)\chi(n),\quad \alpha(n):=\begin{cases}1 & \text{if } n \text{ is a square-full integer,}\\ 0 & \text{otherwise,}\end{cases}\tag{2-9}$$

which is the character sum of square-full integers not exceeding $x$. The next lemma is [Munsch 2014, Lemma 1.3].

**Lemma 2.5.** *Let $p>3$ be an integer, let $\lambda$ be a Dirichlet character mod $p$, and let $\lambda_0$ be the trivial character mod $p$. Then*

$$Q_2(x;\lambda)=\begin{cases}\dfrac{\zeta(\frac{3}{2})}{\zeta(3)}\displaystyle\prod_{p_1\mid p}\Big(\frac{1-p_1^{-1}}{1+p_1^{-\frac{3}{2}}}\Big)x^{\frac{1}{2}}+\dfrac{\zeta(\frac{2}{3})}{\zeta(2)}\prod_{p_1\mid p}\Big(\frac{1-p_1^{-\frac{2}{3}}}{1+p_1^{-1}}\Big)x^{\frac{1}{3}}+O(x^{\frac{1}{6}+\varepsilon}p^{\varepsilon}) & \text{if } \lambda=\lambda_0,\\ \dfrac{L(\frac{3}{2},\chi)}{\zeta(3)}\displaystyle\prod_{p_1\mid p}\Big(1+\frac{1}{p_1}+\frac{1}{p_1^2}\Big)^{-1}x^{\frac{1}{2}}+O(x^{\frac{1}{4}}(\log x)^{\frac{1}{2}}p^{\frac{3}{32}+\varepsilon}) & \text{if } \lambda \text{ is a quadratic character,}\\ \dfrac{L(\frac{2}{3},\chi^2)}{\zeta(2)}\displaystyle\prod_{p_1\mid p}\Big(1+\frac{1}{p_1}\Big)^{-1}x^{\frac{1}{3}}+O(x^{\frac{1}{4}}p^{\frac{1}{4}+\varepsilon}) & \text{if } \lambda \text{ is a cubic character,}\\ O(x^{\frac{1}{4}}(\log x)^{\frac{1}{2}}p^{\frac{11}{32}+\varepsilon}) & \text{if } \lambda^2\ne\lambda_0 \text{ and } \lambda^3\ne\lambda_0.\end{cases}$$

By exactly the same steps of proof as that of [Munsch 2014, p. 562], we extract:

**Lemma 2.6.** *Let $p,q$ be positive integers with $pq>3$, let $\xi^{(j)}$ $(j=1,2)$ be Dirichlet characters mod $pq$, and let $\xi_0$ be the trivial character mod $pq$. If $\xi^{(1)}\ne\xi_0$ and $\xi^{(2)}\ne\xi_0$, then*

$$\sum_{a\le x^{1/2}}\xi^{(1)}(a)\sum_{b\le(x/a^2)^{1/3}}\mu^2(b)\xi^{(2)}(b)\ll x^{\frac{1}{4}}(\log x)^{\frac{1}{2}}(pq)^{\frac{11}{32}+\varepsilon}.$$

## 3. Proof of Theorem 1.1

Keeping the notation as in the statement of Theorem 1.1, let

$$Q_p(x;\ell,q):=\sum_{\substack{n\le x\\ n\equiv\ell \bmod q}}T_2(n)\tag{3-1}$$

denote the number of square-full integers $n\equiv\ell \bmod q$ not exceeding $x$ which are primitive roots mod $p$. By the orthogonality relation for Dirichlet characters mod $q$, we have

$$Q_p(x;\ell,q)=\frac{1}{\phi(q)}\sum_{\chi \bmod q}\bar{\chi}(\ell)\sum_{n\le x}T_2(n)\chi(n).\tag{3-2}$$

Using (1-6), the definition of $T_2(n)$, together with Lemma 2.1 and (2-9), we have

$$Q_p(x;\ell,q)=\frac{\phi(p-1)}{(p-1)\phi(q)}\sum_{d\mid p-1}\frac{\mu(d)}{\phi(d)}\sum_{\chi \bmod q}\bar{\chi}(\ell)\sum_{\lambda\in\Gamma_d}\sum_{n\le x}\alpha(n)\chi(n)\lambda(n),$$

where the sum for $\lambda$ runs over the Dirichlet characters mod $p$ of order $d$. For brevity, let

$$T(x;\chi,\lambda):=\sum_{n\le x}\alpha(n)\chi(n)\lambda(n), \tag{3-3}$$

so that

$$Q_p(x;\ell,q)=\frac{\phi(p-1)}{(p-1)\phi(q)}\sum_{d\mid p-1}\frac{\mu(d)}{\phi(d)}\sum_{\chi \bmod q}\bar{\chi}(\ell)\sum_{\lambda\in\Gamma_d}T(x;\chi,\lambda). \tag{3-4}$$

The Euler product formula for $L(s,\chi)=\sum_{n=1}^{\infty}\chi(n)n^{-s}$ leads to the Dirichlet series of the function $\alpha(n)\chi(n)\lambda(n)$ as

$$\sum_{n=1}^{\infty}\alpha(n)\chi(n)\lambda(n)n^{-s}=\frac{L(2s,\chi^2\lambda^2)L(3s,\chi^3\lambda^3)}{L(6s,\chi^6\lambda^6)}. \tag{3-5}$$

Our task now is to derive asymptotic estimates for $T(x;\chi,\lambda)$, the sum of the product of two characters over square-full integers. There are three main cases. Case 1 deals with $\chi=\chi_0$, the principal character mod $q$. Case 2 deals with $\lambda=\lambda_0$, the principle character mod $p$ and Case 3 is when both $\chi$ and $\lambda$ are nonprincipal characters with respect to their moduli. The subcases run through all possible shapes of the two characters.

Case 1: $\chi=\chi_0$, the principal character mod $q$. If $\lambda$ is any Dirichlet character modulo $p$, then

$$\begin{aligned}
&\sum_{n=1}^{\infty}\alpha(n)\chi_0(n)\lambda(n)n^{-s}\\
&\quad=\frac{L(2s,\chi_0\lambda^2)L(3s,\chi_0\lambda^3)}{L(6s,\chi_0\lambda^6)}\\
&\quad=\prod_{p_1}\left(\frac{(1-\chi_0(p_1)\lambda^2(p_1)p_1^{-2s})(1-\chi_0(p_1)\lambda^3(p_1)p_1^{-3s})}{1-\chi_0(p_1)\lambda^6(p_1)p_1^{-6s}}\right)^{-1}\\
&\quad=\prod_{\substack{p_1\\ p_1\nmid q}}\left(\frac{(1-\lambda^2(p_1)p_1^{-2s})(1-\lambda^3(p_1)p_1^{-3s})}{1-\lambda^6(p_1)p_1^{-6s}}\right)^{-1}\\
&\quad=\prod_{p_1}\left(\frac{(1-\lambda^2(p_1)p_1^{-2s})(1-\lambda^3(p_1)p_1^{-3s})}{1-\lambda^6(p_1)p_1^{-6s}}\right)^{-1}\prod_{p_1\mid q}\frac{(1-\lambda^2(p_1)p_1^{-2s})(1-\lambda^3(p_1)p_1^{-3s})}{1-\lambda^6(p_1)p_1^{-6s}}\\
&\quad=\frac{L(2s,\lambda^2)L(3s,\lambda^3)}{L(6s,\lambda^6)}H_q(s,\lambda)=\left(\sum_{n=1}^{\infty}\alpha(n)\lambda(n)n^{-s}\right)H_q(s,\lambda),
\end{aligned}$$

where the last product runs over primes $p_1\mid q$, and

$$H_q(s,\lambda)=\prod_{p_1\mid q}\frac{1-\lambda^2(p_1)p_1^{-2s}}{1+\lambda^3(p_1)p_1^{-3s}}:=\sum_{n=1}^{\infty}h_q(n,\lambda)n^{-s}. \tag{3-6}$$

Equating coefficients of the Dirichlet series, summing over $n$, and making use of (3-3) and (2-9) leads to

$$T(x:\chi_0,\lambda)=\sum_{n\le x}\alpha(n)\chi_0(n)\lambda(n)=\sum_{n\le x}\sum_{de=n}h_q(d,\lambda)\alpha(e)\lambda(e)$$
$$=\sum_{d\le x}h_q(d,\lambda)\sum_{e\le x/d}\alpha(e)\lambda(e)=\sum_{d\le x}h_q(d,\lambda)Q_2\Big(\frac{x}{d},\lambda\Big). \tag{3-7}$$

We next obtain asymptotic formulas for (3-7) by applying Lemma 2.5 to $Q_2(x/d;\lambda)$.

Subcase 1.1: $\lambda=\lambda_0$, the principal character mod $p$. We get, using also (3-6),

$$\begin{aligned}T(x;\chi_0,\lambda_0)&=\sum_{d\le x}h_q(d,\lambda_0)Q_2\Big(\frac{x}{d},\lambda_0\Big)\\&=\frac{\zeta(\frac32)}{\zeta(3)}\prod_{p_1\mid p}\Big(\frac{1-p_1^{-1}}{1+p_1^{-\frac32}}\Big)x^{\frac12}\sum_{d\le x}\Big(\frac{h_q(d,\lambda_0)}{d}\Big)^{\frac12}\\&\quad+\frac{\zeta(\frac23)}{\zeta(2)}\prod_{p_1\mid p}\Big(\frac{1-p_1^{-\frac23}}{1+p_1^{-1}}\Big)x^{\frac13}\sum_{d\le x}\Big(\frac{h_q(d,\lambda_0)}{d}\Big)^{\frac13}+O\Big(x^{\frac16+\varepsilon}p^{\varepsilon}\sum_{d\le x}\Big(\frac{h_q(d,\lambda_0)}{d}\Big)^{\frac16+\varepsilon}\Big)\\&=\frac{\zeta(\frac32)}{\zeta(3)}\prod_{p_1\mid pq}\Big(\frac{1-p_1^{-1}}{1+p_1^{-\frac32}}\Big)x^{\frac12}+\frac{\zeta(\frac23)}{\zeta(2)}\prod_{p_1\mid pq}\Big(\frac{1-p_1^{-\frac23}}{1+p_1^{-1}}\Big)x^{\frac13}+O(x^{\frac16+\varepsilon}p^{\varepsilon}).\end{aligned}$$

The contribution from $T(x;\chi_0,\lambda_0)$ towards (3-4) is equal to

$$\frac{\phi(p-1)}{\phi(p)\phi(q)}\frac{\mu(1)}{\phi(1)}\bar\chi_0(\ell)T(x;\chi_0,\lambda_0)$$
$$=\frac{\phi(p-1)}{\phi(p)\phi(q)}\Big(\frac{\zeta(\frac32)}{\zeta(3)}\prod_{p_1\mid pq}\Big(\frac{1-p_1^{-1}}{1+p_1^{-\frac32}}\Big)x^{\frac12}+\frac{\zeta(\frac23)}{\zeta(2)}\prod_{p_1\mid pq}\Big(\frac{1-p_1^{-\frac23}}{1+p_1^{-1}}\Big)x^{\frac13}+O(x^{\frac16+\varepsilon}p^{\varepsilon})\Big). \tag{3-8}$$

Subcase 1.2: $\lambda^2=\lambda_0,\ \lambda\ne\lambda_0$, i.e, $\lambda$ a quadratic character mod $p$. Similar reasoning as in the last case leads to

$$T(x;\chi_0,\lambda)=\frac{L(\frac32,\lambda)}{\zeta(3)}\Big(1+\frac1p+\frac{1}{p^2}\Big)^{-1}H_q\big(\tfrac12,\lambda\big)x^{\frac12}+O(x^{\frac14}(\log x)^{\frac12}p^{\frac{3}{32}+\varepsilon}).$$

The contribution from $T(x;\chi_0,\lambda)$ towards (3-4) is equal to

$$\frac{\phi(p-1)}{\phi(p)\phi(q)}\frac{\mu(2)}{\phi(2)}\bar\chi_0(\ell)\sum_{\lambda\in Y_1}T(x;\chi_0,\lambda)$$
$$=-\frac{\phi(p-1)}{\phi(p)\phi(q)}\sum_{\lambda\in Y_1}\Big(\frac{L(\frac32,\lambda)}{\zeta(3)}\Big(1+\frac1p+\frac{1}{p^2}\Big)^{-1}H_q\big(\tfrac12,\lambda\big)x^{\frac12}+O(x^{\frac14}(\log x)^{\frac12}p^{\frac{3}{32}+\varepsilon})\Big). \tag{3-9}$$

Subcase 1.3: $\lambda^3=\lambda_0,\ \lambda\ne\lambda_0$, i.e., $\lambda$ a cubic character mod $p$. As before, similar calculation yields

$$T(x;\chi_0,\lambda)=\frac{L(\frac23,\lambda^2)}{\zeta(2)}\Big(1+\frac1p\Big)^{-1}H_q\big(\tfrac13,\lambda\big)x^{\frac13}+O(x^{\frac14}p^{\frac14+\varepsilon}).$$

The contribution from $T(x;\chi_0,\lambda)$ towards (3-4) is equal to

$$\frac{\phi(p-1)}{\phi(p)\phi(q)}\frac{\mu(3)}{\phi(3)}\bar{\chi}_0(\ell)\sum_{\lambda\in Y_2}T(x;\chi_0,\lambda)$$
$$=-\frac{\phi(p-1)}{2\phi(p)\phi(q)}\sum_{\lambda\in Y_2}\left(\frac{L\left(\frac{2}{3},\lambda^2\right)}{\zeta(2)}\left(1+\frac{1}{p}\right)^{-1}H_q\left(\tfrac{1}{3},\lambda\right)x^{\frac{1}{3}}+O(x^{\frac{1}{4}}p^{\frac{1}{4}+\varepsilon})\right). \quad (3\text{-}10)$$

Subcase 1.4: $\lambda^2\neq\lambda_0,\ \lambda^3\neq\lambda_0$. Similar calculation gives

$$T(x;\chi_0,\lambda)=O(x^{\frac{1}{4}}(\log x)^{\frac{1}{2}}p^{\frac{11}{32}+\varepsilon}).$$

The contribution from $T(x;\chi_0,\lambda)$ towards (3-4) is equal to

$$\frac{\phi(p-1)}{\phi(p)\phi(q)}\sum_{\substack{d\mid p-1\\ d>3}}\frac{\mu(d)}{\phi(d)}\bar{\chi}_0(\ell)\sum_{\lambda\in\Gamma_d\setminus\{Y_1\cup Y_2\}}T(x;\chi_0,\lambda)$$
$$=\frac{\phi(p-1)}{\phi(p)\phi(q)}\sum_{\substack{d\mid p-1\\ d>3}}\frac{\mu(d)}{\phi(d)}\sum_{\lambda\in\Gamma_d\setminus\{Y_1\cup Y_2\}}O(x^{\frac{1}{4}}(\log x)^{\frac{1}{2}}p^{\frac{11}{32}+\varepsilon}). \quad (3\text{-}11)$$

Case 2: $\lambda=\lambda_0$, the principal character mod $p$, and $\chi\neq\chi_0$. From (3-5), steps similar to those at the beginning of Case 1 yield

$$\sum_{n=1}^{\infty}\alpha(n)\chi(n)\lambda_0(n)n^{-s}=\frac{L(2s,\chi^2\lambda_0)L(3s,\chi^3\lambda_0)}{L(6s,\chi^6\lambda_0)}$$
$$=\prod_{p_1}\left(\frac{(1-\chi^2(p_1)\lambda_0(p_1)p_1^{-2s})(1-\chi^3(p_1)\lambda_0(p_1)p_1^{-3s})}{1-\chi^6(p_1)\lambda_0(p_1)p_1^{-6s}}\right)^{-1}$$
$$=\prod_{\substack{p_1\\ p_1\neq p}}\left(\frac{(1-\chi^2(p_1)p_1^{-2s})(1-\chi^3(p_1)p_1^{-3s})}{1-\chi^6(p_1)p_1^{-6s}}\right)^{-1}$$
$$=\frac{(1-\chi^2(p)p^{-2s})(1-\chi^3(p)p^{-3s})}{1-\chi^6(p)p^{-6s}}\prod_{p_1}\left(\frac{(1-\chi^2(p_1)p_1^{-2s})(1-\chi^3(p_1)p_1^{-3s})}{1-\chi^6(p_1)p_1^{-6s}}\right)^{-1}$$
$$=\frac{L(2s,\chi^2)L(3s,\chi^3)}{L(6s,\chi^6)}F_p(s,\chi)=\left(\sum_{n=1}^{\infty}\alpha(n)\chi(n)n^{-s}\right)F_q(s,\chi),$$

where

$$F_p(s,\chi)=\frac{1-\chi^2(p)p^{-2s}}{1+\chi^3(p)p^{-3s}}:=\sum_{n=1}^{\infty}f_p(n,\chi)n^{-s},$$

and consequently,

$$T(x;\chi,\lambda_0)=\sum_{d\leq x}f_p(d,\chi)Q_2\left(\frac{x}{d},\chi\right). \quad (3\text{-}12)$$

Subcase 2.1: $\chi^2=\chi_0,\ \chi\neq\chi_0$. We have

$$T(x;\chi,\lambda_0)=\frac{L(\frac{3}{2},\chi)}{\zeta(3)}\prod_{p_1\mid q}\left(1+\frac{1}{p_1}+\frac{1}{p_1^2}\right)^{-1}F_p(\tfrac{1}{2},\chi)x^{\frac{1}{2}}+O(x^{\frac{1}{4}}(\log x)^{\frac{1}{2}}q^{\frac{3}{32}+\varepsilon}).$$

The contribution from $T(x;\chi,\lambda_0)$ towards (3-4) is equal to

$$\begin{aligned}&\frac{\phi(p-1)}{\phi(p)\phi(q)}\frac{\mu(1)}{\phi(1)}\sum_{\chi\in X_1}\bar{\chi}(\ell)T(x;\chi,\lambda_0)\\&=\frac{\phi(p-1)}{\phi(p)\phi(q)}\sum_{\chi\in X_1}\bar{\chi}(\ell)\left(\frac{L(\frac{3}{2},\chi)}{\zeta(3)}\prod_{p_1\mid q}\left(1+\frac{1}{p_1}+\frac{1}{p_1^2}\right)^{-1}F_p(\tfrac{1}{2},\chi)x^{\frac{1}{2}}+O(x^{\frac{1}{4}}(\log x)^{\frac{1}{2}}q^{\frac{3}{32}+\varepsilon})\right).\end{aligned}\tag{3-13}$$

Subcase 2.2: $\chi^3=\chi_0,\ \chi\neq\chi_0$. We have

$$T(x;\chi,\lambda_0)=\frac{L(\frac{2}{3},\chi^2)}{\zeta(2)}\prod_{p_1\mid q}\left(1+\frac{1}{p_1}\right)^{-1}F_p(\tfrac{1}{3},\chi)x^{\frac{1}{3}}+O(x^{\frac{1}{4}}q^{\frac{1}{4}+\varepsilon}).$$

The contribution from $T(x;\chi,\lambda_0)$ towards (3-4) is equal to

$$\begin{aligned}&\frac{\phi(p-1)}{\phi(p)\phi(q)}\frac{\mu(1)}{\phi(1)}\sum_{\chi\in X_2}\bar{\chi}(\ell)T(x;\chi,\lambda_0)\\&\qquad=\frac{\phi(p-1)}{\phi(p)\phi(q)}\sum_{\chi\in X_2}\bar{\chi}(\ell)\left(\frac{L(\frac{2}{3},\chi^2)}{\zeta(2)}\prod_{p_1\mid q}\left(1+\frac{1}{p_1}\right)^{-1}F_p(\tfrac{1}{3},\chi)x^{\frac{1}{3}}+O(x^{\frac{1}{4}}q^{\frac{1}{4}+\varepsilon})\right).\end{aligned}\tag{3-14}$$

Subcase 2.3: $\chi^2\neq\chi_0,\ \chi^3\neq\chi_0$. We have

$$T(x;\chi,\lambda_0)=O(x^{\frac{1}{4}}(\log x)^{\frac{1}{2}}q^{\frac{11}{32}+\varepsilon}).$$

The contribution from $T(x;\chi,\lambda_0)$ towards (3-4) is equal to

$$\frac{\phi(p-1)}{\phi(p)\phi(q)}\frac{\mu(1)}{\phi(1)}\sum_{\chi\notin X_1\cup X_2}\bar{\chi}(\ell)T(x;\chi,\lambda_0)=\frac{\phi(p-1)}{\phi(p)\phi(q)}\sum_{\chi\notin X_1\cup X_2}\bar{\chi}(\ell)O(x^{\frac{1}{4}}(\log x)^{\frac{1}{2}}q^{\frac{11}{32}+\varepsilon}).\tag{3-15}$$

Case 3: $\chi\neq\chi_0,\ \lambda\neq\lambda_0$, i.e., both are nonprincipal characters.

Subcase 3.1: $\chi^2=\chi_0,\ \lambda^2=\lambda_0$, i.e., both are quadratic characters. We proceed as in [Munsch 2014, p. 560]. Since $n$ is square-full, we can uniquely write $n=a^2b^3$ with $b$ square-free, and so

$$\begin{aligned}T(x;\chi,\lambda)&=\sum_{n\le x}\alpha(n)\chi(n)\lambda(n)=\sum_{a^2b^3\le x}\mu^2(b)\chi(a^2b^3)\lambda(a^2b^3)\\&=\sum_{b\le x^{1/3}}\mu^2(b)\chi(b^3)\lambda(b^3)\sum_{a\le(x/b^3)^{1/2}}\chi(a^2)\lambda(a^2)=\sum_{b\le x^{1/3}}\mu^2(b)\chi(b^3)\lambda(b^3)\sum_{\substack{a\le(x/b^3)^{1/2}\\ \gcd(a,pq)=1}}1.\end{aligned}$$

For any real $H \geq 1$, we have

$$T(x;\chi,\lambda) = \sum_{b\leq H}\mu^2(b)\chi(b^3)\lambda(b^3) \sum_{\substack{a\leq (x/b^3)^{1/2}\\ \gcd(a,pq)=1}} 1 + \sum_{H<b\leq x^{1/3}}\mu^2(b)\chi(b^3)\lambda(b^3) \sum_{\substack{a\leq (x/b^3)^{1/2}\\ \gcd(a,pq)=1}} 1$$
$$= \sum_{b\leq H}\mu^2(b)\chi(b^3)\lambda(b^3) \sum_{\substack{a\leq (x/b^3)^{1/2}\\ \gcd(a,pq)=1}} 1 + \sum_{\substack{a\leq (x/H^3)^{1/2}\\ \gcd(a,pq)=1}} \sum_{H<b\leq (x/a^2)^{1/3}} \mu^2(b)\chi(b^3)\lambda(b^3).$$

The first term is bounded using (2-5), the character shapes and (2-7):

$$\sum_{b\leq H}\mu^2(b)\chi(b^3)\lambda(b^3) \sum_{\substack{a\leq (x/b^3)^{1/2}\\ \gcd(a,pq)=1}} 1$$
$$= \sum_{b\leq H}\mu^2(b)\chi(b)\lambda(b)\left(\frac{\phi(pq)}{pq}\frac{x^{\frac{1}{2}}}{b^{\frac{3}{2}}} + O(\tau(pq))\right)$$
$$= \frac{L\left(\frac{3}{2},\chi\lambda\right)}{\zeta(3)}\prod_{p_1\mid pq}(1-p_1^{-3})^{-1}\frac{\phi(pq)}{pq}x^{\frac{1}{2}} + O(x^{\frac{1}{2}}(pq)^{\frac{3}{16}+\varepsilon}(\log H)H^{-1}) + O(H(pq)^{\varepsilon}).$$

The second term is bounded using the character shapes and the second bound in Lemma 2.3:

$$\sum_{\substack{a\leq (x/H^3)^{1/2}\\ \gcd(a,pq)=1}} \sum_{H<b\leq (x/a^2)^{1/3}} \mu^2(b)\chi(b^3)\lambda(b^3) \ll (pq)^{\frac{3}{16}+\varepsilon}x^{\frac{1}{6}}(\log x) \sum_{a\leq (x/H^3)^{1/2}} a^{-\frac{1}{3}} \ll (pq)^{\frac{3}{16}+\varepsilon}x^{\frac{1}{2}}(\log x)H^{-1}.$$

Choosing $H = x^{1/4}(\log x)^{1/2}(pq)^{3/32}$, we have

$$T(x;\chi,\lambda) = \frac{L\left(\frac{3}{2},\chi\lambda\right)}{\zeta(3)}\prod_{p_1\mid pq}(1-p_1^{-3})^{-1}\frac{\phi(pq)}{pq}x^{\frac{1}{2}} + O(x^{\frac{1}{4}}(\log x)^{\frac{1}{2}}(pq)^{\frac{3}{32}+\varepsilon}).$$

The contribution from $T(x;\chi,\lambda)$ towards (3-4) is equal to

$$\frac{\phi(p-1)}{\phi(p)\phi(q)}\frac{\mu(2)}{\phi(2)}\sum_{\chi\in X_1}\bar{\chi}(\ell)\sum_{\lambda\in Y_1}T(x;\chi,\lambda)$$
$$= -\frac{\phi(p-1)}{\phi(p)\phi(q)}\sum_{\chi\in X_1}\bar{\chi}(\ell)\sum_{\lambda\in Y_1}\left(\frac{L\left(\frac{3}{2},\chi\lambda\right)}{\zeta(3)}\prod_{p_1\mid pq}(1-p_1^{-3})^{-1}\frac{\phi(pq)}{pq}x^{\frac{1}{2}} + O(x^{\frac{1}{4}}(\log x)^{\frac{1}{2}}(pq)^{\frac{3}{32}+\varepsilon})\right). \quad (3\text{-}16)$$

Subcase 3.2: $\chi^3 = \chi_0,\ \lambda^3 = \lambda_0$, i.e., both are cubic characters. We proceed as in [Munsch 2014, p. 561] using the character shapes, for any real $H \geq 1$, to get

$$T(x;\chi,\lambda) = \sum_{n\leq x}\alpha(n)\chi(n)\lambda(n) = \sum_{a^2b^3\leq x}\mu^2(b)\chi(a^2b^3)\lambda(a^2b^3)$$
$$= \sum_{a\leq x^{1/2}}\chi(a^2)\lambda(a^2)\sum_{b\leq (x/a^2)^{1/3}}\mu^2(b)\chi(b)^3\lambda(b)^3 = \sum_{a\leq x^{1/2}}\chi(a^2)\lambda(a^2)\sum_{\substack{b\leq (x/a^2)^{1/3}\\ \gcd(b,pq)=1}}\mu^2(b)$$
$$= \sum_{a\leq H}\chi(a^2)\lambda(a^2)\sum_{\substack{b\leq (x/a^2)^{1/3}\\ \gcd(b,pq)=1}}\mu^2(b) + \sum_{H<a\leq x^{1/2}}\chi(a^2)\lambda(a^2)\sum_{\substack{b\leq (x/a^2)^{1/3}\\ \gcd(b,pq)=1}}\mu^2(b). \quad (3\text{-}17)$$

The first term in (3-17) is bounded by using (2-6) and (2-8):

$$\sum_{a\le H}\chi(a^2)\lambda(a^2)\sum_{\substack{b\le (x/a^2)^{1/3}\\ \gcd(b,pq)=1}}\mu^2(b)$$
$$=\sum_{a\le H}\chi(a^2)\lambda(a^2)\left(\frac{x^{\frac13}}{a^{\frac23}\zeta(2)}\prod_{p_1\mid pq}\left(1+\frac{1}{p_1}\right)^{-1}+O(x^{\frac16}a^{-\frac13}\tau(pq))\right)$$
$$=\frac{x^{\frac12}}{\zeta(2)}\prod_{p_1\mid pq}\left(1+\frac{1}{p_1}\right)^{-1}\sum_{a\le H}\frac{\chi^2(a)\lambda^2(a)}{a^{\frac23}}+O(x^{\frac16}(pq)^{\varepsilon}H^{\frac23})$$
$$=\frac{L\left(\frac23,\chi^2\lambda^2\right)}{\zeta(2)}\prod_{p_1\mid pq}(1+p_1^{-1})^{-1}x^{\frac13}+O(x^{\frac16}(pq)^{\varepsilon}H^{\frac23})+O(x^{\frac13}(pq)^{\frac12}(\log pq)H^{-\frac23}).$$

The second term in (3-17) is bounded by inverting the summation and using Lemma 2.2(I):

$$\sum_{H<a\le x^{\frac12}}\chi(a^2)\lambda(a^2)\sum_{\substack{b\le (x/a^2)^{1/3}\\ \gcd(b,pq)=1}}\mu^2(b)=\sum_{\substack{b\le (x/H^2)^{1/3}\\ \gcd(b,pq)=1}}\mu^2(b)\sum_{H<a\le (x/b^3)^{1/2}}(\chi\lambda)^2(a)$$
$$\ll (pq)^{\frac12}(\log pq)x^{\frac13}H^{-\frac23}.$$

Thus,

$$T(x;\chi,\lambda)=\frac{L\left(\frac23,\chi^2\lambda^2\right)}{\zeta(2)}\prod_{p_1\mid pq}(1+p_1^{-1})^{-1}x^{\frac13}+O(x^{\frac16}(pq)^{\varepsilon}H^{\frac23})+O(x^{\frac13}(pq)^{\frac12}(\log pq)H^{-\frac23}).$$

Choosing $H=x^{1/8}(pq)^{3/8}$, we obtain

$$T(x;\chi,\lambda)=\frac{L\left(\frac23,\chi^2\lambda^2\right)}{\zeta(2)}\prod_{p_1\mid pq}(1+p_1^{-1})^{-1}x^{\frac13}+O(x^{\frac14}(pq)^{\frac14+\varepsilon}).$$

The contribution from $T(x;\chi,\lambda)$ towards (3-4) is equal to

$$\frac{\phi(p-1)}{\phi(p)\phi(q)}\frac{\mu(3)}{\phi(3)}\sum_{\chi\in X_2}\bar\chi(\ell)\sum_{\lambda\in Y_2}T(x;\chi,\lambda)$$
$$=-\frac{\phi(p-1)}{2\phi(p)\phi(q)}\sum_{\chi\in X_2}\bar\chi(\ell)\sum_{\lambda\in Y_2}\left(\frac{L\left(\frac23,\chi^2\lambda^2\right)}{\zeta(2)}\prod_{p_1\mid pq}(1+p_1^{-1})^{-1}x^{\frac13}+O(x^{\frac14}(pq)^{\frac14+\varepsilon})\right). \quad (3\text{-}18)$$

Subcase 3.3: $\chi^2\ne\chi_0$, $\chi^3\ne\chi_0$, $\lambda^2\ne\lambda_0$, $\lambda^3\ne\lambda_0$, i.e., both are nonquadratic and noncubic characters. We proceed as in [Munsch 2014, p. 562]. Similar to the last subcase, using the character shapes, for any real $H\ge 1$, we have

$$T(x;\chi,\lambda)=\sum_{a\le x^{1/2}}\chi(a^2)\lambda(a^2)\sum_{b\le (x/a^2)^{1/3}}\mu^2(b)\chi(b^3)\lambda(b^3)$$
$$=\sum_{a\le H}\chi(a^2)\lambda(a^2)\sum_{b\le (x/a^2)^{1/3}}\mu^2(b)\chi(b^3)\lambda(b^3)+\sum_{H<a\le x^{1/2}}\chi(a^2)\lambda(a^2)\sum_{b\le (x/a^2)^{1/3}}\mu^2(b)\chi(b^3)\lambda(b^3).$$

The first term is bounded using the second bound in Lemma 2.3:

$$\sum_{a\le H}\chi(a^2)\lambda(a^2)\sum_{b\le (x/a^2)^{1/3}}\mu^2(b)(\chi\lambda)^3(b)\ll (pq)^{\frac{3}{16}+\varepsilon}x^{\frac{1}{6}}(\log x)\sum_{a\le H}a^{-\frac{1}{3}}$$
$$\ll (pq)^{\frac{3}{16}+\varepsilon}x^{\frac{1}{6}}(\log x)H^{\frac{2}{3}}.$$

The second term is bounded using Lemma 2.2:

$$\sum_{H<a\le x^{1/2}}\chi(a^2)\lambda(a^2)\sum_{b\le (x/a^2)^{1/3}}\mu^2(b)\chi(b^3)\lambda(b^3)=\sum_{b\le (x/H^2)^{1/3}}\mu^2(b)\chi(b^3)\lambda(b^3)\sum_{H<a\le (x/b^3)^{1/2}}(\chi\lambda)^2(a)$$
$$\ll \sqrt{pq}\log(pq)x^{\frac{1}{3}}H^{-\frac{2}{3}}.$$

Thus,

$$T(x;\chi,\lambda)=O((pq)^{\frac{3}{16}+\varepsilon}x^{\frac{1}{6}}(\log x)H^{\frac{2}{3}})+O((pq)^{\frac{1}{2}}\log(pq)x^{\frac{1}{3}}H^{-\frac{2}{3}}).$$

Choosing $H=x^{1/8}(pq)^{15/64}(\log x)^{-3/4}$, we obtain

$$T(x;\chi,\lambda)=O((pq)^{\frac{11}{32}+\varepsilon}x^{\frac{1}{4}}(\log x)^{\frac{1}{2}}).$$

The contribution from $T(x;\chi,\lambda)$ towards (3-4) is equal to

$$\frac{\phi(p-1)}{\phi(p)\phi(q)}\sum_{\substack{d\mid p-1\\ d>3}}\frac{\mu(d)}{\phi(d)}\sum_{\chi\notin X_1\cup X_2}\bar{\chi}(\ell)\sum_{\lambda\in\Gamma_d\setminus Y_1\cup Y_2}T(x;\chi,\lambda)$$
$$=\frac{\phi(p-1)}{\phi(p)\phi(q)}\sum_{\substack{d\mid p-1\\ d>3}}\frac{\mu(d)}{\phi(d)}\sum_{\chi\notin X_1\cup X_2}\bar{\chi}(\ell)\sum_{\lambda\in\Gamma_d\setminus Y_1\cup Y_2}O((pq)^{\frac{11}{32}+\varepsilon}x^{\frac{1}{4}}(\log x)^{\frac{1}{2}}). \quad (3\text{-}19)$$

Subcase 3.4: $\chi^2=\chi_0$, $\lambda^3=\lambda_0$. Since $\chi$ and $\lambda$ are nonprincipal characters mod $q$ and mod $p$, respectively, with prime $p\nmid q$, in this subcase the product $\chi\lambda$ can be considered as a nonprincipal character mod $pq$. Thus,

$$T(x;\chi,\lambda)=\sum_{a\le x^{1/2}}\chi(a^2)\lambda(a^2)\sum_{b\le (x/a^2)^{1/3}}\mu^2(b)\chi(b^3)\lambda(b^3)$$
$$=\sum_{a\le x^{1/2}}(\chi_0\lambda^2)(a)\sum_{b\le (x/a^2)^{1/3}}\mu^2(b)(\chi^3\lambda_0)(b)=\sum_{a\le x^{1/2}}\xi^{(1)}(a)\sum_{b\le (x/a^2)^{1/3}}\mu^2(b)\xi^{(2)}(b), \quad (3\text{-}20)$$

where $\xi^{(1)}:=\chi_0\lambda^2$ and $\xi^{(2)}:=\chi^3\lambda_0$ are Dirichlet characters modulo $pq$. Since $\xi^{(1)}$ and $\xi^{(2)}$ are nonprincipal characters mod $pq$, by Lemma 2.6, we have

$$T(x;\chi,\lambda)=O((pq)^{\frac{11}{32}+\varepsilon}x^{\frac{1}{4}}(\log x)^{\frac{1}{2}}) \quad (3\text{-}21)$$

and the contribution from $T(x;\chi,\lambda)$ towards (3-4) is equal to

$$\frac{\phi(p-1)}{\phi(p)\phi(q)}\frac{\mu(2)}{\phi(2)}\sum_{\chi\in X_1}\bar{\chi}(\ell)\sum_{\lambda\in Y_2}T(x;\chi,\lambda)=-\frac{\phi(p-1)}{\phi(p)\phi(q)}\sum_{\chi\in X_1}\bar{\chi}(\ell)\sum_{\lambda\in Y_2}O((pq)^{\frac{11}{32}+\varepsilon}x^{\frac{1}{4}}(\log x)^{\frac{1}{2}}). \quad (3\text{-}22)$$

The remaining subcases can all be treated in a manner similar to Subcase 3.4 to yield the same estimate (3-21) for $T(x;\chi,\lambda)$, and their contributions from $T(x;\chi,\lambda)$ towards (3-4) are listed below.

Subcase 3.5: $\chi^2=\chi_0,\ \lambda^2\neq\lambda_0,\ \lambda^3\neq\lambda_0$.

$$\frac{\phi(p-1)}{\phi(p)\phi(q)}\sum_{\substack{d\mid p-1\\ d>3}}\frac{\mu(d)}{\phi(d)}\sum_{\chi\in X_1}\bar{\chi}(\ell)\sum_{\lambda\in\Gamma_d\setminus Y_1\cup Y_2}O((pq)^{\frac{11}{32}+\varepsilon}x^{\frac{1}{4}}(\log x)^{\frac{1}{2}}). \tag{3-23}$$

Subcase 3.6: $\chi^3=\chi_0,\ \lambda^2=\lambda_0,\ \lambda\neq\lambda_0$.

$$\frac{\phi(p-1)}{\phi(p)\phi(q)}\frac{\mu(2)}{\phi(2)}\sum_{\chi\in X_2}\bar{\chi}(\ell)\sum_{\lambda\in Y_1}O((pq)^{\frac{11}{32}+\varepsilon}x^{\frac{1}{4}}(\log x)^{\frac{1}{2}}). \tag{3-24}$$

Subcase 3.7: $\chi^3=\chi_0,\ \lambda^2\neq\lambda_0,\lambda^3\neq\lambda_0$.

$$\frac{\phi(p-1)}{\phi(p)\phi(q)}\sum_{\substack{d\mid p-1\\ d>3}}\frac{\mu(d)}{\phi(d)}\sum_{\chi\in X_2}\bar{\chi}(\ell)\sum_{\lambda\in\Gamma_d\setminus Y_1\cup Y_2}O((pq)^{\frac{11}{32}+\varepsilon}x^{\frac{1}{4}}(\log x)^{\frac{1}{2}}). \tag{3-25}$$

Subcase 3.8: $\chi^2\neq\chi_0,\chi^3\neq\chi_0,\ \lambda^2=\lambda_0,\lambda\neq\lambda_0$.

$$\frac{\phi(p-1)}{\phi(p)\phi(q)}\frac{\mu(2)}{\phi(2)}\sum_{\chi\notin X_1\cup X_2}\bar{\chi}(\ell)\sum_{\lambda\in Y_1}O((pq)^{\frac{11}{32}+\varepsilon}x^{\frac{1}{4}}(\log x)^{\frac{1}{2}}). \tag{3-26}$$

Subcase 3.9: $\chi^2\neq\chi_0,\ \chi^3\neq\chi_0,\ \lambda^3=\lambda_0$.

$$\frac{\phi(p-1)}{\phi(p)\phi(q)}\frac{\mu(3)}{\phi(3)}\sum_{\chi\notin X_1\cup X_2}\bar{\chi}(\ell)\sum_{\lambda\in Y_2}O((pq)^{\frac{11}{32}+\varepsilon}x^{\frac{1}{4}}(\log x)^{\frac{1}{2}}). \tag{3-27}$$

The largest contribution towards the sum in (3-4), i.e., the term containing

$$\frac{\phi(p-1)}{\phi(p)\phi(q)}x^{\frac{1}{2}},$$

comes from (3-8) in Subcase 1.1, (3-9) in Subcase 1.2, (3-13) in Subcase 2.1, and (3-16) in Subcase 3.1, with coefficient equal to $A_{p,q}$ as displayed in (1-7).

The second largest contribution in (3-4), i.e., the term containing

$$\frac{\phi(p-1)}{\phi(p)\phi(q)}x^{\frac{1}{3}},$$

comes from (3-8) in Subcase 1.1, (3-10) in Subcase 1.3, (3-14) in Subcase 2.2, and (3-18) in Subcase 3.2 with coefficient equal to $B_{p,q}$ as displayed in (1-10).

The contribution from the error terms in (3-4) coming from Subcases 1.1–1.4 is, apart from the factor

$$\frac{\phi(p-1)}{\phi(p)\phi(q)},$$

equal to

$$O(x^{\frac{1}{6}+\varepsilon}p^{\varepsilon})-\sum_{\lambda\in Y_1}O(x^{\frac{1}{4}}(\log x)^{\frac{1}{2}}p^{\frac{3}{32}+\varepsilon})-\frac{1}{2}\sum_{\lambda\in Y_2}O(x^{\frac{1}{4}}p^{\frac{1}{4}+\varepsilon})+\sum_{\substack{d\mid p-1\\ d>3}}\frac{\mu(d)}{\phi(d)}\sum_{\lambda\in\Gamma_d\setminus Y_1\cup Y_2}O(x^{\frac{1}{4}}(\log x)^{\frac{1}{2}}p^{\frac{11}{32}+\varepsilon}).$$

The contribution from the error terms in (3-4) coming from Subcases 2.1–2.3 is, apart from the factor

$$\frac{\phi(p-1)}{\phi(p)\phi(q)},$$

equal to

$$\sum_{\chi\in X_1}\bar\chi(\ell)O(x^{\frac14}(\log x)^{\frac12}q^{\frac{3}{32}+\varepsilon})+\sum_{\chi\in X_2}\bar\chi(\ell)O(x^{\frac14}q^{\frac14+\varepsilon})+\sum_{\chi\notin X_1\cup X_2}\bar\chi(\ell)O(x^{\frac14}(\log x)^{\frac12}q^{\frac{11}{32}+\varepsilon}).$$

The contribution from the error terms in (3-4) coming from Subcases 3.1–3.9 is, apart from the factor

$$\frac{\phi(p-1)}{\phi(p)\phi(q)},$$

equal to

$$\begin{aligned}
&-\sum_{\chi\in X_1}\bar\chi(\ell)\sum_{\lambda\in Y_1}O(x^{\frac14}(\log x)^{\frac12}(pq)^{\frac{3}{32}+\varepsilon})\\
&\quad-\frac12\sum_{\chi\in X_2}\bar\chi(\ell)\sum_{\lambda\in Y_2}O(x^{\frac14}(pq)^{\frac14+\varepsilon})\\
&\qquad+\sum_{\substack{d\mid p-1\\ d>3}}\frac{\mu(d)}{\phi(d)}\sum_{\chi\notin X_1\cup X_2}\bar\chi(\ell)\sum_{\lambda\in\Gamma_d\setminus Y_1\cup Y_2}O((pq)^{\frac{11}{32}+\varepsilon}x^{\frac14}(\log x)^{\frac12})\\
&\qquad\quad-\sum_{\chi\in X_1}\bar\chi(\ell)\sum_{\lambda\in Y_2}O((pq)^{\frac{11}{32}+\varepsilon}x^{\frac14}(\log x)^{\frac12})\\
&\qquad\quad+\sum_{\substack{d\mid p-1\\ d>3}}\frac{\mu(d)}{\phi(d)}\sum_{\chi\in X_1}\bar\chi(\ell)\sum_{\lambda\in\Gamma_d\setminus Y_1\cup Y_2}O((pq)^{\frac{11}{32}+\varepsilon}x^{\frac14}(\log x)^{\frac12})\\
&\qquad\quad-\sum_{\chi\in X_2}\bar\chi(\ell)\sum_{\lambda\in Y_1}O((pq)^{\frac{11}{32}+\varepsilon}x^{\frac14}(\log x)^{\frac12})\\
&\qquad\quad+\sum_{\substack{d\mid p-1\\ d>3}}\frac{\mu(d)}{\phi(d)}\sum_{\chi\in X_2}\bar\chi(\ell)\sum_{\lambda\in\Gamma_d\setminus Y_1\cup Y_2}O((pq)^{\frac{11}{32}+\varepsilon}x^{\frac14}(\log x)^{\frac12})\\
&\qquad\quad-\sum_{\chi\notin X_1\cup X_2}\bar\chi(\ell)\sum_{\lambda\in Y_1}O((pq)^{\frac{11}{32}+\varepsilon}x^{\frac14}(\log x)^{\frac12})\\
&\qquad\quad-\frac12\sum_{\chi\notin X_1\cup X_2}\bar\chi(\ell)\sum_{\lambda\in Y_2}O((pq)^{\frac{11}{32}+\varepsilon}x^{\frac14}(\log x)^{\frac12}).
\end{aligned}$$

Taking all the subcases into account, the error term is

$$\begin{aligned}
&\ll(pq)^{\frac{11}{32}+\varepsilon}x^{\frac14}(\log x)^{\frac12}\sum_{\substack{d\mid p-1\\ d>3}}\frac{|\mu(d)|}{\phi(d)}\sum_{\chi \bmod q}|\bar\chi(\ell)|\sum_{\lambda\in\Gamma_d}1\\
&\ll(pq)^{\frac{11}{32}+\varepsilon}x^{\frac14}(\log x)^{\frac12}2^{\omega(p-1)}\phi(q),
\end{aligned}$$

using the estimates

$$\sum_{\lambda\in\Gamma_d}1=O(\phi(d)),\qquad\sum_{\chi \bmod q}|\bar\chi(\ell)|=O(\phi(q)),\qquad\sum_{\substack{d\mid p-1\\ d>3}}|\mu(d)|=O(2^{\omega(p-1)}),$$

and the theorem is proved.

## References

[Bateman and Grosswald 1958] P. T. Bateman and E. Grosswald, "On a theorem of Erdős and Szekeres", *Illinois J. Math.* **2** (1958), 88–98. MR Zbl

[Burgess 1962] D. A. Burgess, "On character sums and primitive roots", *Proc. London Math. Soc.* (3) **12** (1962), 179–192. MR Zbl

[Cai 1997] Y. Cai, "On the distribution of square-full integers", *Acta Math. Sinica* (*N.S.*) **13**:2 (1997), 269–280. MR Zbl

[Cao 1994] X.-D. Cao, "The distribution of square-full integers", *Period. Math. Hungar.* **28**:1 (1994), 43–54. MR Zbl

[Cao 1997] X. Cao, "On the distribution of square-full integers", *Period. Math. Hungar.* **34**:3 (1997), 169–175. MR Zbl

[Chan 2015] T. H. Chan, "Squarefull numbers in arithmetic progression, II", *J. Number Theory* **152** (2015), 90–104. MR Zbl

[Chan and Tsang 2013] T. H. Chan and K. M. Tsang, "Squarefull numbers in arithmetic progressions", *Int. J. Number Theory* **9**:4 (2013), 885–901. MR Zbl

[Erdős and Szekeres 1934] P. Erdős and S. Szekeres, "Über die Anzahl der Abelschen Gruppen gegebener Ordnung und über ein verwandtes zahlentheoretisches Problem", *Acta Sci. Math.* (*Szeged*) **7** (1934), 95–102. Zbl

[Iwaniec and Kowalski 2004] H. Iwaniec and E. Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications **53**, American Mathematical Society, Providence, RI, 2004. MR Zbl

[Liu 1994] H.-Q. Liu, "The distribution of square-full integers", *Ark. Mat.* **32**:2 (1994), 449–454. MR Zbl

[Liu and Zhang 2005] H. Liu and W. Zhang, "On the squarefree and squarefull numbers", *J. Math. Kyoto Univ.* **45**:2 (2005), 247–255. MR Zbl

[Liu and Zhang 2013] H. Liu and T. Zhang, "On the distribution of square-full numbers in arithmetic progressions", *Arch. Math.* (*Basel*) **101**:1 (2013), 53–64. MR Zbl

[Munsch 2014] M. Munsch, "Character sums over squarefree and squarefull numbers", *Arch. Math.* (*Basel*) **102**:6 (2014), 555–563. MR Zbl

[Munsch and Trudgian 2018] M. Munsch and T. Trudgian, "Square-full primitive roots", *Int. J. Number Theory* **14**:4 (2018), 1013–1021. MR Zbl

[Shapiro 1983] H. N. Shapiro, *Introduction to the theory of numbers*, John Wiley & Sons, New York, 1983. MR Zbl

[Srichan 2013] T. Srichan, "Square-full and cube-full numbers in arithmetic progressions", *Šiauliai Math. Semin.* **8** (2013), 223–248. MR Zbl

[Srichan 2020] T. Srichan, "On the distribution of square-full and cube-full primitive roots", *Period. Math. Hungar.* **80**:1 (2020), 103–107. MR Zbl

[Suryanarayana and Sitaramachandra Rao 1973] D. Suryanarayana and R. Sitaramachandra Rao, "The distribution of square-full integers", *Ark. Mat.* **11** (1973), 195–201. MR Zbl

[Wu 1998] J. Wu, "On the distribution of square-full and cube-full integers", *Monatsh. Math.* **126**:4 (1998), 353–367. MR Zbl

[Wu 2001] J. Wu, "On the distribution of square-full integers", *Arch. Math.* (*Basel*) **77**:3 (2001), 233–240. MR Zbl

Received 26 May 2020.

VICHIAN LAOHAKOSOL:

fscivil@ku.ac.th
Department of Mathematics, Faculty of Science, Kasetsart University, Bangkok, Thailand

TEERAPAT SRICHAN:

fscitrp@ku.ac.th
Department of Mathematics, Faculty of Science, Kasetsart University, Bangkok, Thailand

PINTHIRA TANGSUPPHATHAWAT:

t.pinthira@hotmail.com
Department of Mathematics, Faculty of Science and Technology, Phranakorn Rajabhat University, Bangkok, Thailand

# 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. 9 no. 2 2020