Albanian Journal of Mathematics

Për një Shqipëri të shkencës dhe kulturës.

Volume 1, 2007

www.albanian-j-math.com

ALBANIAN JOURNAL
OF MATHEMATICS
Volume 1, Number 1, Pages 3–11
ISSN 1930-1235: (electronic version)

KLEIN-FOUR COVERS OF THE PROJECTIVE LINE IN CHARACTERISTIC TWO

DARREN GLASS

(Communicated by T. Shaska)

Abstract. In this paper we examine curves defined over a field of characteristic 2 which are $(\mathbb{Z}/2\mathbb{Z})^2$-covers of the projective line. In particular, we determine which 2-ranks occur for such curves of a given genus and where possible we give explicit equations for such curves. As a corollary, we show that there exist hyperelliptic curves of genus g and 2-rank σ which contain an additional involution in their automorphism group if and only if $g \equiv \sigma \pmod 2$.

1. Introduction

There are many ways to stratify the moduli space of curves. When working over an algebraically closed field of characteristic $p > 0$, one of the most natural stratifications comes from looking at the p-ranks of the curves. The p-rank of a curve X (or, more precisely, the p-rank of its Jacobian) can be defined as $\dim_{\mathbb{F}_p} \mathrm{Hom}(\mu_p, \mathrm{Jac}(X))$ where μ_p is the kernel of Frobenius on $\mathbb{G}_m$. In particular, curves of p-rank σ will have precisely p^σ distinct p-torsion points on their Jacobian defined over k.

It follows from [3] in characteristic $p > 2$ and [10] in characteristic 2 that there exist curves of each possible 2-rank for every genus. In this note, we investigate what one can say about the 2-ranks of curves which have multiple copies of $\mathbb{Z}/2\mathbb{Z}$ in their automorphism group. More precisely, we consider curves defined over an algebraically closed field of characteristic $p = 2$ which admit an action of $(\mathbb{Z}/2\mathbb{Z})^2$ and such that their quotient by this action is $\mathbb{P}^1$.

In Section 2 of this paper, we introduce notation and recall some results from [3] and [4] about Klein-four covers of the projective line. We also recall some results from the theory of Artin-Schreier covers that will be used to compute the genera and 2-ranks of the relevant curves. Section 3 is concerned with some nonexistence results, and we prove a number of results about the necessary conditions for a given 2-rank to occur. In the fourth section, we prove that the necessary conditions proven in Section 3 are in fact sufficient, and in particular we prove (a stronger version of) the following theorem.

Received by the editors August 7, 2006 and, in revised form, December 13, 2006.
2000 *Mathematics Subject Classification.* Primary: 14Hxx, Secondary: 14H37, 14H45.
Key words and phrases. algebraic curves, automorphism groups, wild ramification.

Theorem 1.1. *Let $g \geq 0$ and $0 \leq \sigma \leq g$. Then there exists a curve X with $G \cong (\mathbb{Z}/2\mathbb{Z})^2 \subseteq \mathrm{Aut}(X)$ and $X/G \cong \mathbb{P}^1$ such that X has genus g and 2-rank σ unless $\sigma = g-1$ or unless g is even and $\sigma = 1$.*

It will follow from the constructions of these curves that they are all defined over the finite field $\mathbb{F}_4$ and in most cases they can be chosen to be defined over $\mathbb{F}_2$. However, it will not always be the case that the 2-torsion points are themselves defined over $\mathbb{F}_4$.

We also relate our results to a result of Zhu in [10] which shows that there exist hyperelliptic curves of every possible 2-rank with no extra automorphisms. The following theorem shows precisely when a hyperelliptic curve can have extra involutions.

Theorem 1.2. *There are hyperelliptic curves of genus g and 2-rank σ which contain an additional involution in their automorphism group if and only if $g \equiv \sigma \pmod 2$.*

Acknowledgements: The author would like to thank R. Pries for many useful conversations.

2. Notation

In this article, we work over an algebraically closed field k of characteristic $p = 2$. We wish to examine curves that are $(\mathbb{Z}/2\mathbb{Z})^2$-covers of the projective line $\mathbb{P}^1_k$. In [3], we examined such curves defined over algebraically closed fields of characteristic $p > 2$ and in particular we used such curves to construct hyperelliptic curves with particular group schemes arising as the p-torsion of their Jacobians. When the characteristic of k is not equal to two, this Hurwitz space of such covers is well-defined (for details, see the results of Wewer in [9]) and in [3] we denoted the moduli space of genus g curves which are $(\mathbb{Z}/2\mathbb{Z})^2$-covers of $\mathbb{P}^1$ by $\mathcal{H}_{g,2}$. However, when the characteristic of k is equal to two we are in the situation of wild ramification, and Wewer's results do not hold. In particular, it is not clear whether $\mathcal{H}_{g,2}$ will be well-defined as a smooth moduli space due to the wild ramification.

From now on, X will be a k-curve of genus g and 2-rank σ which is a $(\mathbb{Z}/2\mathbb{Z})^2$-cover of $\mathbb{P}^1$. Let H_1, H_2, and H_3 be the three subgroups of $(\mathbb{Z}/2\mathbb{Z})^2$ with respect to a fixed basis. Furthermore, let C_1, C_2, and C_3 be the three quotient curves of X by these subgroups. Finally, we define g_i to be the genus of C_i and σ_i to be the 2-rank of C_i. By results of Kani and Rosen in [5], $\mathrm{Jac}(X) \sim \prod \mathrm{Jac}(C_i)$ and therefore it follows that $g_X = g_1 + g_2 + g_3$ and $\sigma_X = \sigma_1 + \sigma_2 + \sigma_3$. We note that X can be viewed as the normalization of the fibre product of any pair of the C_i, and if the C_i are defined over $\mathbb{F}_2$ then X will be defined over $\mathbb{F}_4$. Throughout this paper, we will use α to denote one of the elements of $\mathbb{F}_4$ other than one or zero.

The fact that we have wild ramification restricts some of the information we can learn from this situation, but there is more that we can say. In particular, we know that C_1, C_2, and C_3 must be Artin-Schreier covers, and therefore can be put into the form $C_i : y^2 + y = f_i(x)$ where f_i is a rational function in $xk(x^2)$. In this case, it follows from results of van der Geer and van der Vlugt in [8] that the third quotient is of the form $y^2 + y = f_3(x)$ where $f_3(x) = f_1(x) + f_2(x)$.

Given a cover of curves $X \to Y$, their genera are related by the Riemann-Hurwitz formula (see [7] for details) and if the Galois group is a p-group then their p-ranks are related by the Deuring-Shafarevich formula (see [1] for details). In particular, if the characteristic of k is two and we have a $\mathbb{Z}/2\mathbb{Z}$-cover $X \to Y$ branched at j

points then the Riemann-Hurwitz formula says that the genera of X and Y are related by the formula $g_X = 2g_Y - 1 + \frac{r}{2}$ where r is the degree of the ramification divisor. It follows immediately that $g_Y \leq \frac{g_X+1}{2}$. The Deuring-Shafarevich formula further says that $\sigma_Y = 2\sigma_X - 1 + j$. The following results about the genus and 2-rank of Artin-Schreier curves in characteristic two follow immediately and will be used throughout this note without reference.

Theorem 2.1. *Let $y^2 + y = f(x)$ define a hyperelliptic curve C in characteristic two. Let $f(x)$ have j poles given by $x_1, \ldots, x_j$ and let n_i be the order of the pole at x_i. Without loss of generality we can assume that all of the n_i are odd. Then the genus of C is given by the formula $-1 + \frac{1}{2}\sum(n_i + 1)$ and the 2-rank of C is given by $j - 1$.*

To conclude this introduction we define the $\mathcal{K}$-type of a Klein-four cover $X \to \mathbb{P}^1$ to be the unordered triple $\mathfrak{p} = \{g_1, g_2, g_3\}$ consisting of the genera of the three $\mathbb{Z}/2\mathbb{Z}$ quotients of X. In particular, it follows that the g_i are integers such that $0 \leq g_i \leq \frac{g+1}{2}$ and $g_1 + g_2 + g_3 = g$, so that $\mathfrak{p}$ is a partition of g. We define a partition $\mathfrak{p}$ – and by extension the $\mathcal{K}$-type of a curve – to be unbalanced if it contains an element which is at least $\frac{g}{2}$. In particular, unbalanced partitions are of the form $\{\frac{g}{2}, g_1, g_2\}$ or $\{\frac{g+1}{2}, g_1, g_2\}$ depending on the parity of g. Note that if $0 \in \mathfrak{p}$ it follows immediately that $\mathfrak{p}$ is unbalanced. On the other extreme, a totally balanced partition is when in which all three elements are the same, and therefore $\mathfrak{p} = \{\frac{g}{3}, \frac{g}{3}, \frac{g}{3}\}$.

We note that the $\mathcal{K}$-type of X is technically the type of the cover $X \to \mathbb{P}^1$, and in a small number of cases a curve X can be considered a $(\mathbb{Z}/2\mathbb{Z})^2$-cover of $\mathbb{P}^1$ in more than one way leading to different types. However, we show in [4] that this is rare in characteristic $p \neq 2$ (and happens exactly in the case where $1 \in \mathfrak{p}$). While not stated in that paper, the proof also works in characteristic 2.

3. Nonexistence Results

Throughout this section, X will be $(\mathbb{Z}/2\mathbb{Z})^2$-cover of $\mathbb{P}^1$ with 2-rank equal to σ and with $\mathcal{K}$-type $\mathfrak{p}$. We will give necessary conditions on σ and $\mathfrak{p}$ in order for such a curve X to exist. Recall that a curve is said to be almost-ordinary if it has 2-rank equal to $g - 1$.

Lemma 3.1. *The 2-rank of X cannot equal $g - 1$.*

Proof. Assume X is almost-ordinary. It follows that one of its $\mathbb{Z}/2\mathbb{Z}$ quotients must be almost-ordinary and the other two must be ordinary. Let C_1 and C_2 be the two quotients which are ordinary so that C_1 (resp. C_2) is defined by the equation $y^2 + y = f_1(x)$ (resp. $f_2(x)$) where f_1 (resp. f_2) only has simple poles. Then $f_1 + f_2$ must also have only simple poles and therefore the curve C_3, which is defined by $y^2 + y = f_1(x) + f_2(x)$, must also be ordinary. This gives a contradiction. □

In some cases it happens that a given 2-rank can occur for curves of some $\mathcal{K}$-types but not for curves of other $\mathcal{K}$-types, as the following results indicate.

Lemma 3.2. *If $\sigma = 0$ then $\mathfrak{p} = \{g_1, g_1, g_3\}$ where $g_1 \geq g_3$.*

Proof. Assume X is a curve with 2-rank equal to zero. It follows that all three of the hyperelliptic quotients have 2-rank zero and therefore they can each be defined

by $y^2 + y = f_i(x)$ where each f_i has a single pole at the same point. It follows that (at least) two of these three functions must have a pole of the same order and that the order of the third pole is no larger than these two, and therefore the same statement can be made about the genera.

□

Lemma 3.3. *If $\sigma = 1$ then g is odd and $\mathfrak{p} = \{\frac{g+1}{2}, g_2, g_3\}$ is unbalanced.*

Proof. Assume that σ equals 1. Then two of the hyperelliptic quotients must have 2-rank zero while the third has 2-rank one. It follows without loss of generality that f_1 has a pole of order a at one point and f_2 has a pole of order b at another point where a and b are both odd. In that case we can compute that the curve X is of type $\{\frac{a-1}{2}, \frac{b-1}{2}, \frac{a+b}{2}\}$ which in turn implies that the genus of the curve X is $a + b - 1$ (and is thus odd) while the genus of the curve C_3 is $\frac{a+b}{2} = \frac{g+1}{2}$.

□

A quite different result holds if we look at curves with 2-rank equal to 2.

Lemma 3.4. *If $\sigma = 2$ then $\mathfrak{p} \neq \{g_1, g_1, g_1\}$.*

Proof. Assume X is a curve whose 2-rank is equal to 2. Let C_1, C_2, and C_3 be the three quotient curves and let σ_i be the 2-rank of C_i. Then it follows without loss of generality that either $\sigma_1 = 2$ and $\sigma_2 = \sigma_3 = 0$ or $\sigma_1 = \sigma_2 = 1$ and $\sigma_3 = 0$. However, the first case cannot happen, because it would imply that f_1 would have 3 poles while each of f_2 and f_3 would have a unique pole.

Therefore we must be in the second case, in which f_1 and f_2 each have two poles and f_3 has one pole. We can assume that f_1 and f_2 each have poles at zero which cancel each other out and poles at infinity and that f_3 has a pole only at infinity. Without loss of generality, we may assume that $ord_\infty(f_1) \geq ord_\infty(f_3)$ which will in turn imply that $g_1 > g_3$. Therefore, $\mathfrak{p}$ cannot be a totally balanced partition.

□

Lemma 3.5. *If $\mathfrak{p}$ is unbalanced then $g \equiv \sigma \pmod 2$.*

Proof. If g is odd and $\frac{g+1}{2} \in \mathfrak{p}$ then there exists an involution $\tau \in \mathrm{Aut}(X)$ such that the genus of $C_1 = X/ < \tau >$ is equal to $\frac{g+1}{2}$. It follows from the Riemann-Hurwitz formula that the cover $X \to C_1$ must be étale. Therefore, if we apply the Deuring-Shafarevich formula to $X \to C_1$ we see that $\sigma_X = 2\sigma_{C_1} - 1$ is odd.

Similarly, if g is even and $\frac{g}{2} \in \mathfrak{p}$ then it follows from the Riemann-Hurwitz formula that the cover $X \to C_1$ must be ramified at a single point. Again, it will follow from the Deuring-Shafarevich formula that $\sigma_X = 2\sigma_{C_1}$ must be even.

Therefore, in both cases where we look at curves whose $\mathcal{K}$-types are unbalanced we see that $\sigma_X \equiv g_X \pmod 2$.

□

4. Existence Results

The main result in this section is that the necessary conditions on σ and $\mathfrak{p}$ which were shown in the previous section are also sufficient. In particular, we will prove the following theorem.

Theorem 4.1. *There exist curves of genus g, 2-rank σ, and $\mathcal{K}$-type $\mathfrak{p}$ under the following conditions:*

(1) $\sigma \neq g-1$.
(2) *If* $\sigma = 0$ *then* $\mathfrak{p} = \{g_1, g_1, g_3\}$ *with* $g_3 \leq g_1$.
(3) *If* $\sigma = 1$ *then* $\frac{g+1}{2} \in \mathfrak{p}$.
(4) *If* $\sigma = 2$ *then* $\mathfrak{p} \neq \{g_1, g_1, g_1\}$.
(5) *If* $\mathfrak{p}$ *is unbalanced, then* $g \equiv \sigma \pmod 2$.

We will prove this theorem by induction on σ after looking at some base cases. In particular, we will use the following inductive lemma which says that if Theorem 4.1 holds for σ then it is *almost* immediate that it will hold for $\sigma + 3$.

Lemma 4.2. *If there exists a curve* X *of genus* g, 2*-rank* σ *and* $\mathcal{K}$*-type* $\mathfrak{p} = \{g_1, g_2, g_3\}$ *then there exists a curve* $\tilde{X}$ *of genus* $g+3$ *and* 2*-rank* $\sigma+3$ *which has* $\mathcal{K}$*-type* $\hat{\mathfrak{p}} = \{g_1+1, g_2+1, g_3+1\}$.

Proof. Assume that the three hyperelliptic quotients of X are defined by the equations $y^2 + y = f_i(x)$, where without loss of generality we may assume that none of the f_i have poles at infinity. Then we define $\tilde{f}_1 = f_1 + x$, $\tilde{f}_2 = f_2 + \alpha x$ and $\tilde{f}_3 = f_3 + (\alpha+1)x$ where α is one of the elements of $\mathbb{F}_4$ other than one or zero. It is clear that $\tilde{f}_3 = \tilde{f}_1 + \tilde{f}_2$ and that the curve $\tilde{X}$ defined by the fibre product of $y^2 + y = \tilde{f}_1(x)$ and $y^2 + y = \tilde{f}_2(x)$ will have the desired properties.

□

We begin proving the necessary base cases by showing that Theorem 4.1 is true for small values of σ.

Lemma 4.3. *Let* $\mathfrak{p} = \{g_1, g_1, g_3\}$ *with* $g_3 \leq g_1$. *Then there exist curves of* $\mathcal{K}$*-type* $\mathfrak{p}$ *and* 2*-rank* $\sigma = 0$.

Proof. Let $a = 2g_1 + 1$ and $b = 2g_3 + 1$ and define $f_1 = x^a$ and $f_3 = x^b$ so that $f_2 = f_1 + f_3 = x^a + x^b$. Then the curves defined by $y^2 + y = f_i(x)$ all have 2-rank equal to zero, and the genera of the curves $y^2 + y = f_1(x)$ and $y^2 + y = f_2(x)$ will each be g_1 while the genus of the curve defined by $y^2 + y = f_3(x)$ will be g_3. Our construction now shows that the relevant fibre product will have the desired properties. □

Lemma 4.4. *Let* g *be odd and let* $\mathfrak{p}$ *be an unbalanced partition (ie* $\frac{g+1}{2} \in \mathfrak{p}$*). Then there are curves* X *of genus* g *and* $\mathcal{K}$*-type* $\mathfrak{p}$ *with* 2*-rank equal to one.*

Proof. Let $\mathfrak{p} = \{\frac{g+1}{2}, g_2, g_3\}$ and define $a = 2g_2 + 1$ and $b = 2g_3 + 1$. The curve C_2 defined by $y^2 + y = x^a$ will have genus g_2 and 2-rank equal to 0 and the curve C_3 defined by $y^2 + y = \frac{1}{x^b}$ will have genus g_3 and 2-rank equal 0. If we look at the fibre product of C_2 and C_3, it will be a $(\mathbb{Z}/2\mathbb{Z})^2$-cover of $\mathbb{P}^1$ and the third hyperelliptic quotient C_1 will be defined by the equation $y^2 + y = x^a + \frac{1}{x^b}$. In particular, the genus of C_1 will be $g_1 = \frac{a+b}{2} = g_2 + g_3 + 1 = \frac{g+1}{2}$ where $g = g_1 + g_2 + g_3$ is the genus of the fibre product C. Similarly, we see that the 2-rank of C is equal to one as desired.

□

Lemma 4.5. *Let* $\mathfrak{p}$ *be a partition which is neither completely balanced or, if* g *is odd, unbalanced. Then there exist curves of* $\mathcal{K}$*-type* $\mathfrak{p}$ *and* 2*-rank equal to two.*

Proof. Let $\mathfrak{p} = \{g_1, g_2, g_3\}$ with $g_1 \geq g_2 \geq g_3$. Let $a = 2g_3 + 1$, $b = 2(g_1 - g_3) - 1$ and $c = 2(g_2 + g_3 - g_1) + 1$. It is clear that a, b, and c are all odd, and that

$a \geq c$. Furthermore, $b \geq 1$ because $\mathfrak{p}$ is not completely balanced and $c \geq 1$ because $g_1 \leq g/2$. If we now let $f_1 = x^a + \frac{1}{x^b}$ and $f_2 = \alpha x^c + \frac{1}{x^b}$ we see that $f_3 = f_1 + f_2 = x^a + \alpha x^c$ and a simple computation shows that the fibre product X will have 2-rank equal to 2 and $\mathcal{K}$-type $\mathfrak{p}$.

□

Note that in all of the above situations, the case of $\sigma = g-1$ is eliminated. Next, we show that the necessary condition on the 2-ranks of curves with unbalanced $\mathcal{K}$-types from Theorem 3.5 is actually sufficient.

Lemma 4.6. *For any unbalanced partition $\mathfrak{p}$, there will be curves of $\mathcal{K}$-type $\mathfrak{p}$ and 2-rank σ as long as $g \equiv \sigma \pmod 2$.*

Proof. In order to prove this lemma we must show that there are curves of $\mathcal{K}$-type $\mathfrak{p} = \{\frac{g}{2}, g_1, g_2\}$ for all even 2-ranks and curves of $\mathcal{K}$-type $\mathfrak{p} = \{\frac{g+1}{2}, g_1, g_2\}$ for all odd σ.

We first consider the case when g is odd and $\mathfrak{p}$ is unbalanced, so that $\mathfrak{p} = \{\frac{g+1}{2}, g_1, g_2\}$ with $g_1 \geq g_2$. We note that we can construct hyperelliptic covers $C_1 \to \mathbb{P}^1$ and $C_2 \to \mathbb{P}^1$ so that the genus of C_i is g_i and the 2-rank of C_i is k_i for all $0 \leq k_i \leq g_i$. Furthermore, after modifying $C \to \mathbb{P}^1$ by a projective linear transformation of $\mathbb{P}^1$, one can assume that the branch loci of the two covers are distinct. If we let X be the fibre product of C_1 and C_2 and consider the third hyperelliptic quotient of X we see that it will have genus $g_1 + g_2 + 1$ and 2-rank $k_1 + k_2 + 1$. If we choose k_1 and k_2 so that $k_1 + k_2 = k$ then X will have 2-rank equal to σ and $\mathcal{K}$-type $\mathfrak{p}$.

Next, we will construct a curve with 2-rank equal to $2m$ and $\mathcal{K}$-type $\{\frac{g}{2}, \frac{g}{2}, 0\}$. We first note that we can find a hyperelliptic curve of genus $\frac{g}{2}$ with 2-rank equal to m for $0 \leq m \leq \frac{g}{2}$. Let us assume that this curve C_1 is defined by the equation $y^2 + y = f_1(x)$ where f_1 has a pole at infinity. Let f_2 be some constant multiple of x so that $f_3 = f_1 + f_2$ will have the same poles (with the same orders) as f_1. Note that if the order of the pole of f_1 at ∞ is greater than one then we can choose this constant multiple to simply be x. If $ord_\infty(f_1) = 1$ then we need to choose a multiple so that f_3 still has a pole at infinity, but we are guaranteed a choice of this multiple defined over $\mathbb{F}_4$. It follows from our construction that the curve X will have 2-rank $2k$ and the desired $\mathcal{K}$-type.

Finally, we consider the case where $\mathfrak{p} = \{\frac{g}{2}, g_1, g_2\}$ with g_1 and g_2 both positive and we wish to show that there will be curves of all even 2-ranks. We note that $\hat{\mathfrak{p}} = \{\frac{g}{2} - 1, g_1 - 1, g_2 - 1\}$ gives an unbalanced partition of $g - 3$. We may assume that $\sigma \geq 4$ (the case $\sigma = 0$ was handled in Lemma 4.3 and the case of $\sigma = 2$ was taken care of by Lemma 4.5), so $\sigma - 3$ will be a positive odd number. In particular, the above argument shows that there exists a curve $\hat{X}$ of $\mathcal{K}$-type $\hat{\mathfrak{p}}$ and 2-rank $\sigma - 3$. The conclusion now follows from Lemma 4.2.

□

For all $\sigma \geq 3$, Theorem 4.1 says that there are no restrictions other than this parity condition for unbalanced $\mathcal{K}$-types and the case where $\sigma = g - 1$. We now show this concretely for $\sigma = 3, 4$, and 5.

Lemma 4.7. *If $\sigma = 3$ then Theorem 4.1 holds. In particular, there are curves of 2-rank equal to three of all $\mathcal{K}$-types except the case where g is even and $\frac{g}{2} \in \mathfrak{p}$.*

Proof. If g is odd and $\mathfrak{p}$ is unbalanced then the result follows from Lemma 4.6. Therefore, it suffices to consider the case where $\mathfrak{p} = \{g_1, g_2, g_3\}$ with $1 \leq g_3 \leq g_2 \leq g_1 \leq \frac{g-1}{2}$. Set $a = 2(g_1 - g_3) + 1$, $b = 2g_3 - 1$ and $d = 2(g_2 + g_3 - g_1) - 1$. We note that our hypotheses imply that a, b, and d are all odd positive numbers with $b \geq d$. Now, let $f_1 = x^a + \frac{\alpha}{x^b}$, $f_2 = x^a + \frac{1}{x^d}$ and $f_3 = f_1 + f_2$. Then the curve defined by $y^2 + y = f_i$ will have genus g_i and 2-rank equal to one, and therefore X will be a curve of 2-rank equal to three.

□

Lemma 4.8. *If $\sigma = 4$ then Theorem 4.1 holds. In particular, there are curves of 2-rank equal to four of all $\mathcal{K}$-types except the case where $g = 5$ or when g is odd and $\frac{g+1}{2} \in \mathfrak{p}$.*

Proof. Assume that $\mathfrak{p} = \{g_1, g_2, g_3\}$ where $g_1 > g_2 \geq g_3$. Let $a = 2g_2 - 1$, $b = 2(g_1 - g_2) - 1$, and $c = 2(g_2 + g_3 - g_1) + 1$. One can easily check that a, b, and c are all positive odd numbers as the fact that $\frac{g+1}{2} \notin \mathfrak{p}$ implies that $g_1 \leq g_2 + g_3$. Furthermore, we see that $a \geq c$. Let $f_1 = x^a + \frac{1}{x^b} + \frac{1}{x+1}$, $f_3 = x^c + \frac{1}{x^b}$, and $f_2 = f_1 + f_3$. Then the curve defined by the equation $y^2 + y = f_i(x)$ has genus g_i and the fibre product X will have genus g and 2-rank $\sigma = 4$ as desired.

On the other hand, assume that $g_1 = g_2 \geq g_3 \geq 2$. In this case, let $a = 2g_1 - 1$ and $b = 2g_3 - 3$. Then it is clear that a and b are positive odd integers with $a > b$. If we define $f_1 = x^a + \frac{1}{x}$ and $f_3 = x^b + \frac{1}{x} + \frac{1}{x+1}$ we can see that the curves will have the desired properties.

For the partition $\mathfrak{p} = \{\frac{g}{2}, \frac{g}{2}, 0\}$ the lemma follows from Lemma 4.6, so it suffices to consider the case where g is odd and $\mathfrak{p} = \{\frac{g-1}{2}, \frac{g-1}{2}, 1\}$. We note that $g \neq 5$, so we may assume that $g \geq 7$. Let $f_1 = x^{g-4} + \frac{1}{x} + \frac{1}{x+1}$ and $f_2 = \alpha x$. These equations define curves with the desired genera and 2-ranks.

□

Lemma 4.9. *If $\sigma = 5$ then Theorem 4.1 holds. In particular, if $g \geq 7$ there are curves of 2-rank equal to five of all $\mathcal{K}$-types except the case where $g = 6$ or the case where g is even and $\frac{g}{2} \in \mathfrak{p}$.*

Proof. Let $\mathfrak{p} = \{g_1, g_2, g_3\}$ be a partition of g with $0 \leq g_3 \leq g_2 \leq g_1 \leq \frac{g+1}{2}$. We wish to show that there are curves of $\mathcal{K}$-type $\mathfrak{p}$ and 2-rank equal to five unless $g_1 = \frac{g}{2}$ (in which case g will be even). If $g_1 = \frac{g+1}{2}$ then the result follows from Lemma 4.6.

If $g_1 \leq \frac{g-1}{2}$ then it follows that $g_3 > 0$ and thus $\hat{\mathfrak{p}} = \{g_1 - 1, g_2 - 1, g_3 - 1\}$ gives a partition of $g - 3$ all of whose entries are at most $\frac{g-3}{2}$. Thus, by Lemma 4.5 there are curves of $\mathcal{K}$-type $\hat{p}$ of 2-rank equal to 2 unless $\hat{p}$ (and therefore $\mathfrak{p}$) is completely balanced. By the induction argument in Lemma 4.2 we therefore have curves whose 2-rank is equal to five in of $\mathcal{K}$-type $\mathfrak{p}$.

It remains to consider the case where $\mathfrak{p}$ is totally balanced: that is, where $g_1 = g_2 = g_3 = a > 2$. To deal with this case, let $f_1 = x^a + \frac{1}{x^a}$ and $f_2 = x^a + \frac{1}{(x-1)^{a-2}} + \frac{1}{x-\alpha}$ and $f_3 = f_1 + f_2$. One can easily compute that these choices will lead to a curve X of $\mathcal{K}$-type $\{a, a, a\}$ whose 2-rank is equal to 5.

□

Before proving the main theorem, there is one more base case that we need to consider.

Lemma 4.10. *Let g be odd and $\frac{g-1}{2} \in \mathfrak{p}$ but $\frac{g+1}{2} \notin \mathfrak{p}$. Then there are curves in of $\mathcal{K}$-type $\mathfrak{p}$ with 2-rank equal to $2m$ for all $0 \leq m \leq \frac{g-3}{2}$.*

Proof. Let $\mathfrak{p} = \{\frac{g-1}{2}, g_1, g_2\}$ with $g_1 \geq g_2 > 0$ and let $\sigma = 2m$ be as above. Because $\sigma \leq g-3$ we have that $m \leq g_1 + g_2 - 2$ and therefore we can choose m_1 and m_2 so that $m_1 + m_2 = m$ but $m_i < g_i$. In particular, we can define a function $h_1(x)$ which has m_1 poles (none of which are at infinity) so that the curve C_1 defined by $y^2 + y = x^3 + h_1(x)$ will have genus g_1 and 2-rank m_1. Similarly, we can choose h_2 with poles distinct from those of h_1 so that the curve C_2 defined by $y^2 + y = \alpha x^3 + h_2(x)$ will have genus g_2 and 2-rank m_2.

If we look at the normalization of the fibre product of C_1 and C_2 we see that the third quotient will be defined by the equation $y^2 + y = (\alpha + 1)x^3 + h_1(x) + h_2(x)$ and therefore will have genus $g_1 + g_2 - 1$ and 2-rank $m_1 + m_2 = m$. Thus, the curve X has $\mathcal{K}$-type $\{\frac{g-1}{2}, g_1, g_2\}$ and has 2-rank equal to $2m$, as desired.

□

We are finally ready to prove Theorem 4.1.

Proof. Given the results of the above lemmata, it suffices to consider the case where $\sigma \geq 6$. In this case, we only need to prove that there are curves of 2-rank equal to σ in every partition if $g \equiv \sigma \pmod 2$ and that there are curves of 2 rank equal to σ in every partition whose entries are all at most $\frac{g-1}{2}$ if $g \not\equiv \sigma \pmod 2$.

If $0 \in \mathfrak{p}$ then $\mathfrak{p}$ must be unbalanced, and therefore we only need to consider the case where $g \equiv \sigma \pmod 2$. The result then follows from Lemma 4.6. Similarly, if $\frac{g+1}{2} \in \mathfrak{p}$ the result follows from Lemma 4.6.

If $\mathfrak{p} = \{\frac{g-1}{2}, g_1, g_2\}$ then it follows from Lemma 4.10 that there are curves of every even 2-rank strictly less than $g-1$ of $\mathcal{K}$-type $\mathfrak{p}$. To construct the curves of odd 2-rank σ, we note that g_1 and g_2 must be positive, and therefore $\hat{\mathfrak{p}} = \{\frac{g-3}{2}, g_1 - 1, g_2 - 1\}$ gives an unbalanced partition of $g-3$. Furthermore, $g - 3 \equiv \sigma - 3 \pmod 2$ and therefore there are curves of $\mathcal{K}$-type $\hat{\mathfrak{p}}$ of 2-rank equal to $\sigma - 3$ by Lemma 4.6. The result then follows from the inductive process described in Lemma 4.2.

If all entries of $\mathfrak{p}$ are at least 1 and at most $\frac{g-2}{2}$, we note $\hat{\mathfrak{p}} = \{g_1 - 1, g_2 - 1, g_3 - 1\}$ gives a partition of $\hat{g} = g - 3$ such that each $\hat{g_i} = g_i - 1$ is at most $\frac{\hat{g}-1}{2}$ and therefore there exist curves of 2-rank $\sigma - 3$ and $\mathcal{K}$-type $\hat{\mathfrak{p}}$. By the inductive procedure described in Lemma 4.2 we can construct a curve of $\mathcal{K}$-type $\mathfrak{p}$ with 2-rank equal to σ, proving the theorem.

□

In [10], Zhu proves that there exist hyperelliptic curves with no extra automorphisms of every possible 2-rank. The following result shows that, depending on the 2-rank, there may or may not be hyperelliptic curves that *do* admit an extra involution.

Corollary 4.11. *There are hyperelliptic curves of genus g and 2-rank σ which contain an additional involution in their automorphism group if and only if $g \equiv \sigma \pmod 2$.*

Proof. It is well known that the hyperelliptic involution is contained in the center of the automorphism group of a curve (see [6] for one proof in characteristic two). Therefore, if there is another involution in the automorphism group then we must have a Klein-four action on the curve and therefore we will be in the setup above.

Furthermore, it follows that the partition $\mathfrak{p}$ corresponding to this curve contains a zero and is therefore either $\mathfrak{p} = \{\frac{g+1}{2}, \frac{g-1}{2}, 0\}$ or $\mathfrak{p} = \{\frac{g}{2}, \frac{g}{2}, 0\}$. In either case, the partition is unbalanced and therefore $g \equiv \sigma \pmod 2$ by Theorem 3.5.

Conversely, it follows from Theorem 4.1 that if $g \equiv \sigma \pmod 2$ then there will exist curves in this partition, which will therefore be both hyperelliptic and contain an extra involution.

□

We note that this does not answer the question of the automorphism groups fully, as the curves may have automorphisms of degree greater than two. We examine the question of the possible 2-ranks of hyperelliptic curves with extra automorphisms in depth in [2].

References

[1] R. M. Crew. Etale p-covers in characteristic p. *Compositio Math.*, 52(1):31–45, 1984.

[2] D. Glass. The 2-ranks of hyperelliptic curves with extra automorphisms. *preprint*, 2006.

[3] D. Glass and R. Pries. Hyperelliptic curves with prescribed p-torsion. *Manuscripta Math.*, 117(3):299–317, 2005.

[4] D. Glass and R. Pries. On the moduli space of Klein four covers of the projective line. In *Computational aspects of algebraic curves*, volume 13 of *Lecture Notes Ser. Comput.*, pages 58–70. World Sci. Publ., Hackensack, NJ, 2005.

[5] E. Kani and M. Rosen. Idempotent relations and factors of Jacobians. *Math. Ann.*, 284(2):307–327, 1989.

[6] C. Lehr and M. Matignon. Automorphism groups for p-cyclic covers of the affine line. *Compos. Math.*, 141(5):1213–1237, 2005.

[7] J.-P. Serre. *Local fields*, volume 67 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1979. Translated from the French by Marvin Jay Greenberg.

[8] G. van der Geer and M. van der Vlugt. Fibre products of artin-schrier curves and generalized hamming weights of codes. *Journal of Combinatorial Theory, Series A*, 1995.

[9] S. Wewers. Construction of Hurwitz spaces. Thesis.

[10] H. J. Zhu. Hyperelliptic curves over $\mathbb{F}_2$ of every 2-rank without extra automorphisms. *Proc. Amer. Math. Soc.*, 134(2):323–331, 2006.

Department of Mathematics, Gettysburg College, Gettysburg, PA 17325

E-mail address: dglass@gettysburg.edu

ALBANIAN JOURNAL
OF MATHEMATICS
Volume 1, Number 1, Pages 13–16
ISSN 1930-1235(electronic version)

THE SKOLEM PROBLEM FOR 2×2 MATRICES, ARCTANGENTS AND RECURSIVE SOLVABILITY

JODY M. LOCKHART

(Communicated by D. Joyner)

Abstract. A new short proof of the solvability of the Skolem problem for two by two matrices is given.

1. Introduction

Decision problems for groups have been studied very extensively starting in the 1950's. More recently, decision problems for finite sets of matrices have been investigated. In 1970, Paterson [5] showed that the mortality problem is unsolvable; a set of $n \times n$ matrices is said to be *mortal* if some finite product of elements in the set is the zero matrix and the mortality problem is the problem of deciding if finite sets of $n \times n$ matrices with integer entries are mortal.

The problem that we consider is the Skolem problem. The Skolem problem is the problem of deciding for a given square matrix with integer entries whether there is some positive power of the matrix that has zero as its entry in the upper right corner. In 1997, V. Halava [3] gave a proof of the solvability of the Skolem problem for 2×2 matrices. The Skolem problem for 3×3 and 4×4 matrices was solved in 1985 by N. K. Vereshchagin [7] and the problem for 5×5 matrices was solved in 2005 by V. Halava, T. Harju, M. Hirvensalo, and J. Karhumaki [4]. In this note, we give a short new proof for the 2×2 case that ties the Skolem problem to a problem about arctangents and uses a beautiful result of J. H. Conway, C. Radin, and L. Sadun [2] about geodetic angles.

2. Skolem and Arctangent Problems

In this section, we first show that the Skolem problem for 2×2 matrices with integer entries and with real eigenvalues is solvable. Then we show that the Skolem problem for 2×2 matrices with integer entries and with non-real eigenvalues reduces to the following problem about arctangents.

Arctangent Problem: Given positive integers m and n, is $\arctan(\frac{\sqrt{m}}{n})$ a rational multiple of π?

Received by the editors December 1, 2006 and, in revised form, February 10, 2007.
2000 *Mathematics Subject Classification.* Primary 15A36.
Key words and phrases. Skolem problem, arctangents.

Proposition 1. *A 2×2 matrix A of integers with real eigenvalues has a positive integer power with upper right corner zero if and only if either A or A^2 has upper right corner zero.*

Proof. Let $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ with $a_{ij} \in \mathbb{Z}$ for $i, j = 1, 2$. If A has two distinct real eigenvalues λ_1 and λ_2, then there is a nonsingular matrix $B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$ such that $B^{-1}AB = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$. Then, for any positive integer n,

$$\begin{aligned} A^n &= B \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}^n B^{-1} \\ &= B \begin{bmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{bmatrix} B^{-1} \\ &= \begin{bmatrix} * & \frac{b_{11}b_{12}}{d}(\lambda_2^n - \lambda_1^n) \\ * & * \end{bmatrix}, \end{aligned} \tag{1}$$

where d is the determinant of B. If $b_{11}b_{12} = 0$, then the upper right corner of A^1 is zero and we are done. Otherwise, the upper right corner of A^n is zero if and only if $\lambda_1^n = \lambda_2^n$. Since λ_1 and λ_2 are real numbers, $\lambda_1^n = \lambda_2^n$ for some positive integer n if and only if $\lambda_1 = \pm\lambda_2$ if and only if $\lambda_1^2 = \lambda_2^2$ if and only if the upper right corner of A^2 is zero.

Next, suppose that A has one real eigenvalue λ. Then there is a nonsingular matrix $B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$ such that either

$$B^{-1}AB = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \quad \textit{or} \quad B^{-1}AB = \begin{bmatrix} \lambda & 0 \\ 1 & \lambda \end{bmatrix}.$$

If $B^{-1}AB = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$, then $A = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$ and A^1 has upper right corner zero. If $B^{-1}AB = \begin{bmatrix} \lambda & 0 \\ 1 & \lambda \end{bmatrix}$, then

$$\begin{aligned} A^n &= B \begin{bmatrix} \lambda & 0 \\ 1 & \lambda \end{bmatrix}^n B^{-1} \\ &= B \begin{bmatrix} \lambda^n & 0 \\ n\lambda^{n-1} & \lambda^n \end{bmatrix} B^{-1} \\ &= \frac{1}{d} \begin{bmatrix} * & -b_{12}^2 n\lambda^{n-1} \\ * & * \end{bmatrix}, \end{aligned} \tag{2}$$

and the upper right corner of A^n is zero if and only if $b_{12} = 0$ or $\lambda = 0$. If $b_{12} = 0$ then the upper right corner of A^1 is zero, and if $\lambda = 0$ then the upper right corner of A^2 is zero.

□

Corollary 2. *The Skolem problem for 2×2 matrices with integer coefficients and with real eigenvalues is solvable.*

Next, consider the case of non-real eigenvalues.

Proposition 3. *The Skolem problem for 2×2 matrices with integer coefficients and with non-real eigenvalues reduces to the arctangent problem.*

Proof. Let A be a 2×2 matrix as above and let λ_1 and λ_2 be its eigenvalues. Since λ_1 and λ_2 are non-real conjugates, they are unequal and as in Eq. (1) of the proposition above, we get

$$A^n = \begin{bmatrix} * & \frac{b_{11}b_{12}}{d}(\lambda_2^n - \lambda_1^n) \\ * & * \end{bmatrix},$$

where $B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$ and $B^{-1}AB = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$. As above, if $b_{11}b_{12} = 0$, the upper right corner of A^1 is 0. If $b_{11}b_{12} \neq 0$, then the upper right corner of A^n is zero if and only if $\lambda_1^n = \lambda_2^n$. Therefore, the Skolem problem reduces to the problem of determining the existence of a positive integer n such that $\lambda_1^n = \lambda_2^n$.

Computing the eigenvalues of A, we get

$$\begin{aligned} \lambda_j &= \frac{a_{11} + a_{22} \pm \sqrt{(a_{11} + a_{22})^2 - 4(a_{11}a_{22} - a_{21}a_{12})}}{2} \text{ , for } j = 1, 2 \\ &= \frac{\ell \pm i\sqrt{m}}{2}, \text{ where } \ell, m \in \mathbb{Z} \text{ and } m > 0. \end{aligned}$$

Let $\theta = \arctan(\frac{\sqrt{m}}{\ell})$ and $r = \frac{1}{2}\sqrt{\ell^2 + m}$. Since $\lambda_1 = r\, e^{i\theta}$ and $\lambda_2 = r\, e^{-i\theta}$,

$$\lambda_1^n = \lambda_2^n \Leftrightarrow e^{in\theta} = e^{-in\theta} \Leftrightarrow e^{2in\theta} = 1 \Leftrightarrow n\theta = \pi t,$$

for some $t \in \mathbb{Z}$. So there is a positive integer n such that the upper right corner of A^n is zero if and only if there is a positive integer n and an integer $t \neq 0 (\mathrm{mod}\, n)$ such that $\theta = \frac{t}{n}\pi$. Therefore, there is a positive integer n such that the upper right corner of A^n is zero if and only if $\arctan(\frac{\sqrt{m}}{\ell}) \in \pi(\mathbb{Q} - \mathbb{Z})$. Since we know that $\arctan(\frac{\sqrt{m}}{\ell}) \notin \pi\mathbb{Z}$, the problem reduces to that of deciding whether or not $\arctan(\frac{\sqrt{m}}{\ell}) \in \pi\mathbb{Q}$.

□

3. Solvability of Arctangent and Skolem Problems

In this section, we will see that the arctangent problem is solvable and thus the Skolem problem for 2×2 matrices over $\mathbb{Z}$ is also solvable. That the arctangent problem is solvable follows from a result of J. H. Conway, C. Radin, and L. Sadun [2]. In [2], they define a *pure geodetic angle* θ to be an angle such that "any one (and therefore each) of its six squared trigonometric functions is rational (or infinite)." The angles $\arctan(\frac{\sqrt{m}}{\ell})$ that we are interested in are pure geodetic angles. Since $\arctan(-\alpha) = -\arctan\alpha$, we may assume that ℓ is positive. Conway et al [2] define angles $\langle p \rangle_d$ satisfying the following conditions.

Condition 1: **Theorem** (Conway, Radin and Sadun) Every pure geodetic angle is uniquely expressible as a rational multiple of π plus an integral linear combination of the angles $\langle p \rangle_d$.

Condition 2: Rewrite $\frac{\sqrt{m}}{\ell}$ as $\frac{b\sqrt{d}}{a}$ where $a, b, d \in \mathbb{Z}^+$, $\gcd(a, b) = 1$, and d is square free. Then the only $\langle p \rangle_d$ that occur in the expression for $\arctan(\frac{b\sqrt{d}}{a})$ are those for which p is a prime divisor of $a^2 + db^2$ and for which the ideal (p) splits in

$\mathcal{O}_d$, the ring of integers of $\mathbb{Q}(\sqrt{-d})$.

If any combination of the $\langle p \rangle_d$ were a rational multiple of π, the expression would not be unique. Therefore, θ is a rational multiple of π if and only if no $\langle p \rangle_d$ occurs in its expression. So, to determine whether $\arctan(\frac{\sqrt{m}}{\ell})$ is a rational multiple of π, we need to determine whether its expression contains any $\langle p \rangle_d$.

Recall that an ideal I is *prime* if $xy \in I$ implies that $x \in I$ or $y \in I$. By definition, an ideal (p) *splits* if $(p) = P_1 P_2 ... P_k$, where the P_i are distinct prime ideals and $k > 1$. An ideal (p) *ramifies* if the factorization of (p) contains a repeated prime ideal. Thus, an ideal splits if and only if it is not prime and is unramified. It is known that if p is an odd prime, then (p) ramifies in $\mathcal{O}_d$ if and only if p divides d ([6], p. 101) and (p) is prime in $\mathcal{O}_d$ if and only if $-d$ is not a nonzero square modulo p ([2], Theorem 5, p. 329). For $p = 2$, the ideal (p) splits in $\mathcal{O}_d$ if and only if $d \equiv 7 (\mathrm{mod}\, 8)$ ([2], Theorem 6, p. 329) . Therefore, we have the following results.

Theorem 4. *Let a, b, and d be positive integers for which $\gcd(a, b) = 1$ and d is square free. Then $\arctan(\frac{b\sqrt{d}}{a})$ is a rational multiple of π if and only if all prime factors p of $a^2 + db^2$ satisfy the following conditions.*

(i) If p is an odd prime then either $p|d$ or $-d$ is not a nonzero square modulo p.
(ii) If $p = 2$, then $d \neq 7 (\mathrm{mod}\, 8)$.

Corollary 5. *There is an algorithm which, given positive integers ℓ and m determines whether or not $\arctan(\frac{\sqrt{m}}{\ell})$ is a rational multiple of π.*

Corollary 6. *The Skolem problem for 2×2 matrices of integers is solvable.*

References

[1] J. Cassaigne and J. Karhumaki, Examples of undecidable problems for 2-generator matrix semigroups, *Theoretical Computer Science, 204 (1998), 29 - 34.*

[2] J. H. Conway, C. Radin, and L. Sadun, On angles whose squared trigonometric functions are rational, *Discrete and Computational Geometry, 22 (1999), 321 - 332.*

[3] V. Halava, Decidable and Undecidable Problems in Matrix Theory, *TUCS Technical Report No. 127 (1997).*

[4] V. Halava, T. Harju, M. Hirvensalo, and J. Karhumaki, Skolem's Problem - On the Border Between Decidability and Undecidability, *TUCS Technical Report No. 683 (2005).*

[5] M. S. Paterson, Unsolvability in 3×3 matrices, *Studies in App. Math. 49 (1970)*, 105 - 107.

[6] H. Pollard, *The Theory of Algebraic Numbers,* Carus Mathematical Monographs, 9, Mathematical Association of America, Washington, DC, 1950.

[7] N. K. Vereshchagin, The problem of the appearance of a zero in a linear recursive sequence, *Mat. Zametki 38, no. 2*, pp. 177 - 189, 347 (1985). English translation - Occurrence of zero in a linear recursive sequence, *Math. Notes 38, nos 1-2*, pp. 609-615, (1985).

Department of Mathematics,, U. S. Naval Academy,, Annapolis,Maryland
E-mail address: jml@usna.edu

ALBANIAN JOURNAL
OF MATHEMATICS
Volume 1, Number 1, Pages 17–30
ISSN 1930-1235: (electronic version)

THE CRITERIA OF RIESZ, HARDY-LITTLEWOOD ET AL. FOR THE RIEMANN HYPOTHESIS REVISITED USING SIMILAR FUNCTIONS

STEFANO BELTRAMINELLI AND DANILO MERLINI

Abstract. The original criteria of Riesz and of Hardy-Littlewood concerning the truth of the Riemann Hypothesis (RH) are revisited and further investigated in light of the recent formulations and results of Maslanka and of Baez-Duarte concerning a representation of the Riemann Zeta function. Then we introduce a general set of similar functions with the emergence of Poisson-like distributions and we present some numerical experiments which indicate that the RH may barely be true.

1. Introduction

It is well known that there are many different criteria for the truth of the Riemann Hypothesis (RH). Some of these are not directly related to the important high level computations and developments concerning the non trivial zeros of the Riemann Zeta function. In fact, at the beginning of the century M. Riesz, and later G.H. Hardy and J.E. Littlewood (among other important results in number theory) found a criterion of "classical type" for the truth of the RH. The above criteria are related to some series involving values of the Zeta function outside the critical strip, i.e. at integers arguments of the Zeta function [8, 10], and in a numerical context, very accurate calculations are needed toward a "possible kind of verification" of the RH.

In the literature important remarks have been given by leading mathematicians (see for example, those cited in [4]). We may think that such criteria may have a limited interest since, with them, one should work outside the critical strip. It is, in fact, true that in dealing with the above criteria one needs the use of arguments of the Zeta function outside the critical strip, and problems of interchange of summations are present. As an example, in the above criteria, if one uses the formula established by the authors, one should give a meaning to an integration over the real line, which exists only for finite intervals. In order to obtain finite numerical results which give "satisfactory" values to the functions supposed to be equal to the reciprocal of the Zeta function outside and inside the critical strip, the integration should be carried out using a special sequence of upper limit of integration extending to infinity [7].

But lately, there have been new developments and rigorous results in connection with this kind of problem: first a "regularization" of the representations of the

Received by the editors August 11, 2006 and, in revised form, November 10, 2006.
2000 *Mathematics Subject Classification.* Primary 11M26.
Key words and phrases. Riemann Zeta function, Riemann Hypothesis, Criteria of Riesz, Hardy-Littlewood and Baez-Duarte.

Zeta function (a pioneering work by Maslanka [9]), followed (in particular) by a new rigorous discrete formulation with theorems concerning the above criteria (the works of Baez-Duarte [1, 2, 3, 4]).

In light of these new approaches, we thought that some of the above criteria deserved still more study, at least in the direction of some numerical experiments. Thus, we introduce additional functions containing two parameters, in order to have further confidence in the numerical results of the experiments.

The content of this work is as follows: in Section 2 we define a general set of functions with two parameters α and β in the spirit of Riesz and of Hardy-Littlewood and then obtain the discrete "representation" of the reciprocal of the Zeta function of our set by means of the two parameter Pochammer's polynomials with their coefficients. For the reader the discussion of the conditions are then given in Appendix A and in Appendix B (they follow strictly the ingenious method of Baez-Duarte for the Riesz case $\alpha = \beta = 2$). In Section 3 we then obtain in some "limit", a Poisson distribution for the coefficients c_k of the Pochammer's polynomials; this is useful in the context of the numerical experiments. These are presented in Section 4 where many various limiting cases are treated. In the case of increasing values of the parameter β, the experiments indicate that the Poisson distribution becomes more and more exact and the sequence c_k becomes a constant which can be evaluated.

We may argue that in the context of the range of validity of the experiments we present the RH may barely be true.

2. The model

We now consider a set of functions with two parameters ($\alpha > 1, \beta > 0$) to obtain $\frac{1}{\zeta(s)}$. These are simply an extension of these two cases: the first (with $\alpha = \beta = 2$) introduced and studied by Riesz [10], the second one (where $\alpha = 1$ and $\beta = 2$) by Hardy-Littlewood [8].

Let $\mu(n)$ be the Möbius function of argument n, where:

$$\mu(n) = \begin{cases} 1, \text{ if } n = 1 \\ (-1)^k, \text{ if } n \text{ is a product of } k \text{ distinct primes} \\ 0, \text{ if } n \text{ contains a square} \end{cases}$$

Let $s = \sigma + it$ be a complex variable. For $\Re(s) > \rho = 1$ one has $\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}$.

Following the original idea of Riesz and Hardy-Littlewood, we now introduce the two-parameters family of functions given by:

$$\varphi(s;\alpha,\beta) := \frac{1}{\Gamma(-\frac{s-\alpha}{\beta})} \int_0^\infty \sum_{n=1}^{\infty} \frac{\mu(n)}{n^\alpha} e^{-\frac{x}{n^\beta}} x^{-\left(\frac{s-\alpha}{\beta}+1\right)} dx \tag{1}$$

so that expanding the right-hand side in powers of x, we obtain:

$$\begin{aligned} \varphi(s;\alpha,\beta) &= \frac{1}{\Gamma(-\frac{s-\alpha}{\beta})} \int_0^\infty \sum_{n=1}^{\infty} \frac{\mu(n)}{n^\alpha} \sum_{k=0}^{\infty} \frac{(-1)^k x^k}{k! n^{\beta k}} x^{-\left(\frac{s-\alpha}{\beta}+1\right)} dx \\ &= \frac{1}{\Gamma(-\frac{s-\alpha}{\beta})} \int_0^\infty \psi(x;\alpha,\beta) x^{-\left(\frac{s-\alpha}{\beta}+1\right)} dx \end{aligned}$$

where

$$\psi(x;\alpha,\beta)=\sum_{k=0}^{\infty}\frac{(-1)^k x^k}{k!}\frac{1}{\zeta(\alpha+\beta k)} \tag{2}$$

The function $\psi(x;\alpha,\beta)$ was introduced by Riesz (case $\psi(x;2,2)$) and by Hardy-Littlewood (case $\psi(x;1,2)$).

If $\psi(x;\alpha,\beta)\sim\frac{A}{x^{\frac{\alpha-\rho}{\beta}-\epsilon}}$ for some ϵ and for large x, then

$$|\varphi(s;\alpha,\beta)|\leq\left|\frac{1}{\Gamma(-\frac{s-\alpha}{\beta})}\right|\int_0^{\infty}\frac{A}{x^{\frac{\alpha-\rho}{\beta}+\frac{\Re(s)-\alpha}{\beta}+1-\epsilon}}dx\leq\left|\frac{1}{\Gamma(-\frac{s-\alpha}{\beta})}\right|\int_0^{\infty}\frac{A}{x^{1+\frac{\Re(s)-\rho}{\beta}-\epsilon}}dx$$

would exist and would eventually be given by $\frac{1}{\zeta(s)}$ with $\zeta(s)\neq 0$ if we choose $\Re(s)>\rho+\beta\epsilon$.

Let $\rho=\frac{1}{2}$. For $\alpha=\beta=2$ we have:

$$\psi(x;2,2)\sim\frac{A}{x^{3/4-\epsilon}}$$

and for $\alpha=1,\beta=2$:

$$\psi(x;1,2)\sim\frac{A}{x^{1/4-\epsilon}}$$

On the other hand expanding (1) in a similar way, we have that:

$$\begin{aligned}\varphi(s;\alpha,\beta)&=\frac{1}{\Gamma(-\frac{s-\alpha}{\beta})}\int_0^{\infty}\sum_{n=1}^{\infty}\frac{\mu(n)}{n^{\alpha}}e^{x(1-\frac{1}{n^{\beta}})}e^{-x}x^{-\left(\frac{s-\alpha}{\beta}+1\right)}dx\\&=\frac{1}{\Gamma(-\frac{s-\alpha}{\beta})}\sum_{k=0}^{\infty}\sum_{n=1}^{\infty}\frac{\mu(n)}{n^{\alpha}}\left(1-\frac{1}{n^{\beta}}\right)^k\int_0^{\infty}\frac{1}{k!}x^{k-\frac{s-\alpha}{\beta}-1}e^{-x}dx\\&=\frac{1}{\Gamma(-\frac{s-\alpha}{\beta})}\sum_{k=0}^{\infty}\sum_{n=1}^{\infty}\frac{\mu(n)}{n^{\alpha}}\left(1-\frac{1}{n^{\beta}}\right)^k\frac{1}{k!}\Gamma(k-\frac{s-\alpha}{\beta})\\&=\frac{1}{\Gamma(-\frac{s-\alpha}{\beta})}\sum_{k=0}^{\infty}c_k\prod_{r=1}^{k}\left(1-\frac{\frac{s-\alpha}{\beta}+1}{r}\right)\Gamma(-\frac{s-\alpha}{\beta})\end{aligned}$$

Thus:

$$\varphi(s;\alpha,\beta)=\sum_{k=0}^{\infty}c_k P_k(\frac{s-\alpha}{\beta}+1) \tag{3}$$

where $P_k(x):=\prod_{r=1}^{k}(1-\frac{x}{r})$ are the Pochhammer polynomials and the sequences:

$$c_k(\alpha,\beta):=\sum_{n=1}^{\infty}\frac{\mu(n)}{n^{\alpha}}\left(1-\frac{1}{n^{\beta}}\right)^k \tag{4}$$

were already studied by Baez-Duarte [2, 3] in the special case $\alpha=\beta=2$. For another sequence appearing in an expansion of ζ, still for $\alpha=\beta=2$ see the work of Maslanka [9].

Let $\Re(s)>\rho+\epsilon$ ($\epsilon>0$ and $\rho\in[1,\infty[$). From a theorem of Baez-Duarte [2, 3], which says that $|P_k(s)|\leq A\cdot k^{-\Re(s)}$ where A is a constant depending on $|s|$, for

large values of k we have that:

$$|\varphi(s;\alpha,\beta)| \leq \mathrm{A}\sum_{k=0}^{\infty} |c_k|\, k^{-\left(\frac{\rho+\epsilon-\alpha}{\beta}+1\right)} \tag{5}$$

In Appendix A we show that if $\alpha > 1$ and $\beta > 0$ the following holds unconditionally:

$$q_k \ll \frac{1}{k^{\frac{\alpha-1}{\beta}}}$$

where

$$q_k = \sum_{n=1}^{\infty} \frac{1}{n^{\alpha}}\left(1-\frac{1}{n^{\beta}}\right)^k$$

Then we obtain:

$$|\varphi(s;\alpha,\beta)| \leq \sum_{k=0}^{\infty} \frac{1}{k^{\frac{\alpha-1}{\beta}}} \cdot \frac{A}{k^{\frac{\rho+\epsilon-\alpha}{\beta}+1}} \leq \sum_{k=0}^{\infty} \frac{A}{k^{\frac{\rho+\epsilon-1}{\beta}+1}} \leq A\sum_{k=0}^{\infty} \frac{1}{k^{1+\frac{\epsilon}{\beta}}} < \infty$$

From this it follows that we can interchange integration and summation in the earlier calculations of φ and thus for $\Re(s) > 1$ we obtain (6) below, i.e. a representation of $[\zeta(s)]^{-1}$:

$$\varphi(s;\alpha,\beta) = \frac{1}{\zeta(s)} = \sum_{k=0}^{\infty} c_k P_k(\frac{s-\alpha}{\beta}+1), \quad \Re(s) > 1 \tag{6}$$

Now for $\Re(s) > \rho + \epsilon$ ($\epsilon > 0$ and $\rho \in [\frac{1}{2}, \infty[$), still from the theorem of Baez-Duarte [2, 3] i.e. that

$$\left|P_k(\frac{s-\alpha}{\beta}+1)\right| \leq \frac{A}{k^{\frac{\Re(s)-\alpha}{\beta}+1}}$$

and assuming:

$$|c_k| \ll \frac{B}{k^{\frac{1}{\beta}(\alpha-\rho-\epsilon)}}$$

then the above series given by (3) converges uniformely. In fact for $\Re(s) > \rho + \epsilon$ we have:

$$|\varphi(s;\alpha,\beta)| \leq \sum_{k=0}^{\infty} \frac{B}{k^{\frac{1}{\beta}(\alpha-\rho-\epsilon)}} \frac{A}{k^{\frac{\Re(s)-\alpha}{\beta}+1}} = \sum_{k=0}^{\infty} \frac{C}{k^{1+\frac{1}{\beta}(\Re(s)-\rho-\epsilon)}}$$

Following Baez-Duarte the series $\varphi(s;\alpha,\beta)$ extends analytically $\frac{1}{\zeta(s)}$ to the half plane $\Re(s) > \rho = \frac{1}{2}$.

We have thus obtained for our family of sequences with parameters α, β that a necessary and sufficient condition for $\zeta(s) \neq 0$ in the half plane $\Re(s) > \rho$ ($\rho \in [\frac{1}{2}, \infty[$) is given by:

$$|c_k(\alpha,\beta)| \leq \frac{\text{const}}{k^{\frac{1}{\beta}(\alpha-\rho-\epsilon)}} \qquad \forall\epsilon > 0, \forall\alpha > 1, \forall\beta > 0 \tag{7}$$

The necessity of the condition (7) is proved in Appendix B.

Remark 1. Instead of using the Möbius function μ in c_k, one may use (for the numerical computations) the formula involving values of the Zeta function:

$$
\begin{aligned}
c_k &= \sum_{n=1}^{\infty} \frac{\mu(n)}{n^{\alpha}} \left(1 - \frac{1}{n^{\beta}}\right)^k \\
&= \sum_{n=1}^{\infty} \frac{\mu(n)}{n^{\alpha}} \sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{1}{n^{\beta j}} = \sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{1}{\zeta(\alpha + \beta j)}
\end{aligned} \tag{8}
$$

Remark 2. From the bound above it follows not only theoretically but also in the context of a numerical analysis that it will be equally difficult to treat the case $\rho \in [\frac{1}{2}, 1]$, for example $\rho = \frac{3}{4}$, as will be the case $\rho = \frac{1}{2} + \epsilon$ with ϵ small. Below in Section 4 we will also treat the case $\alpha = \frac{7}{2}$.

Remark 3. The condition for the truth of the RH using Riesz and Hardy-Littlewood functions $\psi(x)$ is essentially the same as the one using the discrete function c_k with $k \in \mathbb{N}$. In a previous work [7] independent of the present one (which essentially uses the Baez-Duarte idea and theorems) some numerical results were obtained for $\psi(x)$ in the case of the Hardy-Littlewood function ($\alpha = 1, \beta = 2$) by integration in the x-space. The discrete version using the function c_k of the discrete variable k [2, 3, 9] has advantages in the numerical computations which will be presented below. Before this we present another way to control the sequence c_k in a numerical context.

3. Poisson like distribution

We still consider the sequence c_k given by:

$$
c_k = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^{\alpha}} \left(1 - \frac{1}{n^{\beta}}\right)^k
$$

Then,

$$
\begin{aligned}
c_k &= \sum_{n=1}^{\infty} \frac{\mu(n)}{n^{\alpha}} e^{k \ln(1 - \frac{1}{n^{\beta}})} \\
&= \sum_{n=1}^{\infty} \frac{\mu(n)}{n^{\alpha}} e^{-\frac{k}{n^{\beta}}} e^{k(\ln(1 - \frac{1}{n^{\beta}}) + \frac{1}{n^{\beta}})} \\
&= \sum_{n=1}^{\infty} \frac{\mu(n)}{n^{\alpha}} e^{-\frac{k}{n^{\beta}}} e^{\Delta(k,n,\beta)}
\end{aligned}
$$

Notice that $\Delta < 0$. For β large we set $\Delta = 0$ to obtain the following approximation:

$$
c_k \cong \sum_{n=1}^{\infty} \frac{\mu(n)}{n^{\alpha}} e^{-\frac{k}{n^{\beta}}}
$$

With this approximation we see that c_k becomes equal to $\psi(x = k)$ of (2) as may easily be checked. Moreover:

$$
c_k \cong \sum_{n=1}^{\infty} \frac{\mu(n)}{n^{\alpha}} e^{k(1 - \frac{1}{n^{\beta}})} e^{-k} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^{\alpha}} \sum_{p=0}^{\infty} \frac{k^p}{p!} \left(1 - \frac{1}{n^{\beta}}\right)^p e^{-k}
$$

Thus:

$$c_k \cong \sum_{p=0}^{\infty} c_p \frac{k^p}{p!} e^{-k} \tag{9}$$

We are in the presence of a Poisson distribution for the c_p: in this way, in our numerical computations, we may control in a "more satisfactory" way the values of c_k. The approximation for c_k by means of the Poisson distribution for the c_k we found, will be more satisfactory with increasing values of β and for large values of k. We may also use the approximation given by (9) in which the upper limit of summation will be given by N instead of ∞, i.e. for large k,

$$c_k \cong \sum_{p=0}^{N} c_p \frac{k^p}{p!} e^{-k} \tag{10}$$

4. Numerical experiments

4.1. **The case $\alpha = \frac{7}{2}$ and $\beta = 4$.** This is a case of interest since the behaviour of the c_k at large values of k is expected to be the same as the case $\alpha = \beta = 2$ [2, 3, 10]. In fact from (7) we ask that for $\Re(s) > \frac{1}{2}$:

$$|c_k(7/2, 4)| \leq \frac{C}{k^{\frac{7/2-1/2-\epsilon}{4}}} \sim \frac{k^{\frac{\epsilon}{4}}}{k^{\frac{3}{4}}} \sim |c_k(2, 2)| \tag{11}$$

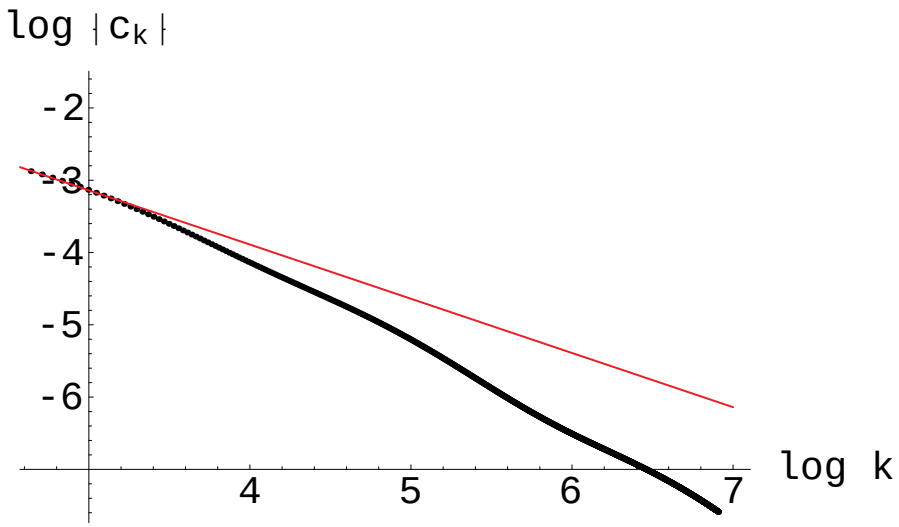


Figure 1. Plot of $\log |c_k| = C - \frac{3}{4} \log k$ together with the straight line of slope$-\frac{3}{4}$.

As a first illustration of the behaviour of c_k (even if k is small) we give in the Figures 1, 2 and 3 the plot respectively of $\log |c_k|, \log(|c_k \log k|)$ and $\log(|c_k (\log k)^2|)$ as a function of $\log k$ for k up to 1000 together with the straight line with slope $-\frac{3}{4}$

which is tangent to the curves at some point. The c_k were computed calculating (4) until $n = 10000$.

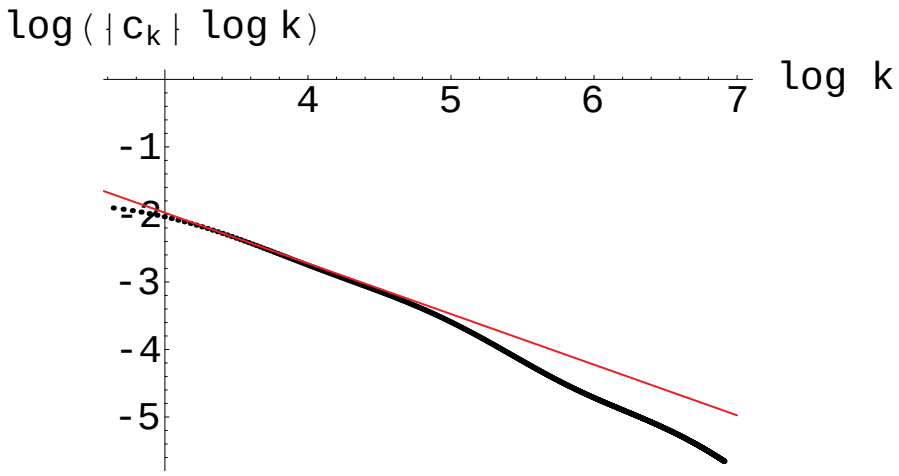


FIGURE 2. Plot of $\log(|c_k| \log k) = C - \frac{3}{4} \log k$ together with the tangent straight line of slope $-\frac{3}{4}$.

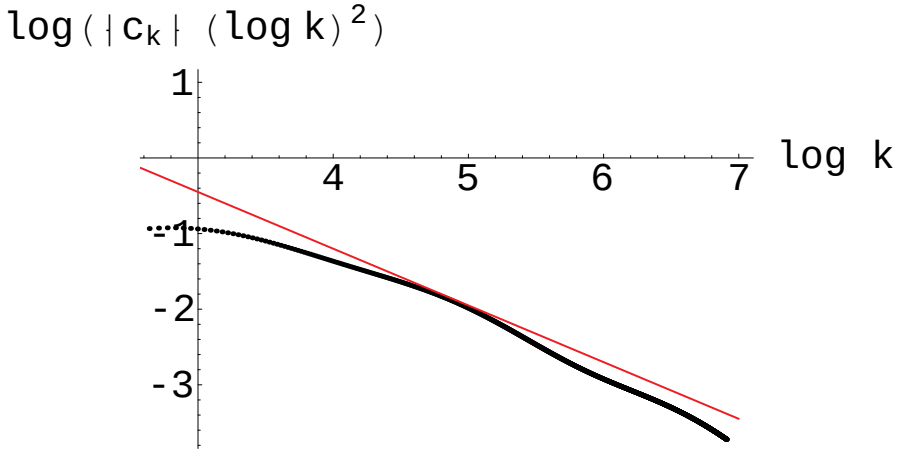


FIGURE 3. Plot of $\log(|c_k|(\log k)^2) = C - \frac{3}{4} \log k$ together with the tangent straight line of slope $-\frac{3}{4}$.

This experiment indicates that c_k, for k up to 1000, may decay more fast than $\frac{C}{(\log k)^2 k^{\frac{3}{4}}}$ as announced by Baez-Duarte in [2] for the case $\alpha = \beta = 2$, i.e. more fast then the bound (11) if the RH is true (see the necessary condition in Appendix B), this of course in the above range of k. For bigger values of k see the Footnote and [3].

4.2. **The case $\alpha = \frac{7}{2}$ with $\beta \to \infty$.** Let $\alpha = \frac{7}{2}$ be fixed, from (8) as β increases we get:

$$\begin{aligned}\lim_{\beta\to\infty} c_k &= \lim_{\beta\to\infty} \sum_{j=0}^{k} (-1)^j \tbinom{k}{j} \tfrac{1}{\zeta(7/2+\beta j)} = \tbinom{k}{0} \tfrac{1}{\zeta(7/2)} - 1 + 1 + \sum_{j=1}^{k} \tbinom{k}{j}(-1)^j \\ &= \tfrac{1}{\zeta(7/2)} - 1 + \sum_{j=0}^{k} \tbinom{k}{j}(-1)^j\end{aligned}$$

Thus:

$$\lim_{\beta\to\infty} c_k = \frac{1}{\zeta(7/2)} - 1 \cong -0.112479 \qquad \forall k \in \mathbb{N} \tag{12}$$

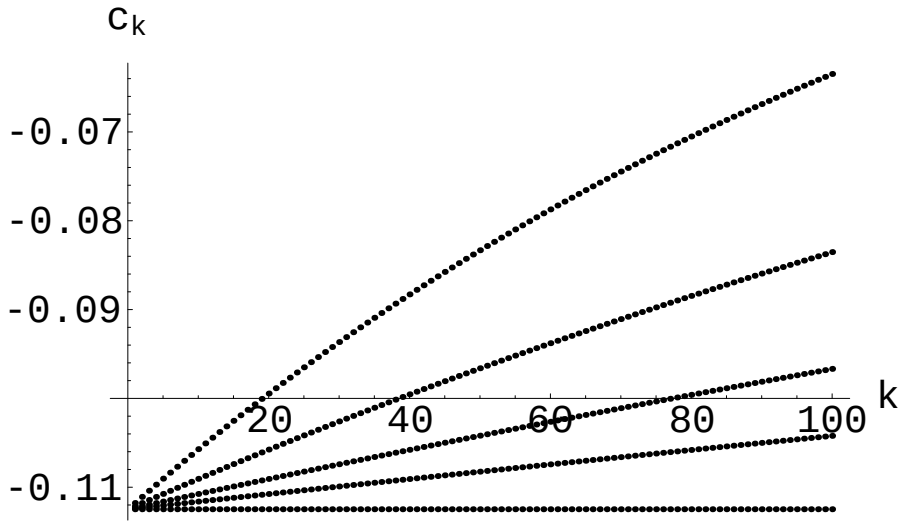


FIGURE 4. Plot of c_k for $\alpha = 7/2$ and $\beta = 4, 5, 6, 7, 20$ (from top to bottom).

Our numerical experiments convalidate these results. We calculated the first 100 c_k for $\beta = 4, 5, 6, 7, 20$. For $\beta = 20$ we get already a convergence to the theoretical limit (12), see Figure 4. So, this infinite β limit obtained by the numerical calculations for low values of k (up to 100) indicates that RH may barely be true (see (7) as $\beta \to \infty$).

4.3. **The Poisson distribution.** To demonstrate the goodness of the approximation's formula (10) we computed the c_k until $k = 1000$ for the case $\alpha = \frac{7}{2}, \beta = 4$ (using (4)). Then using these already computed c_k we calculated also the first 500 c_k of (10). We plotted these two curves together. In Figure 5 we see that from $k \cong 40$ the Poisson approximation is essentially the same as the real sequence.

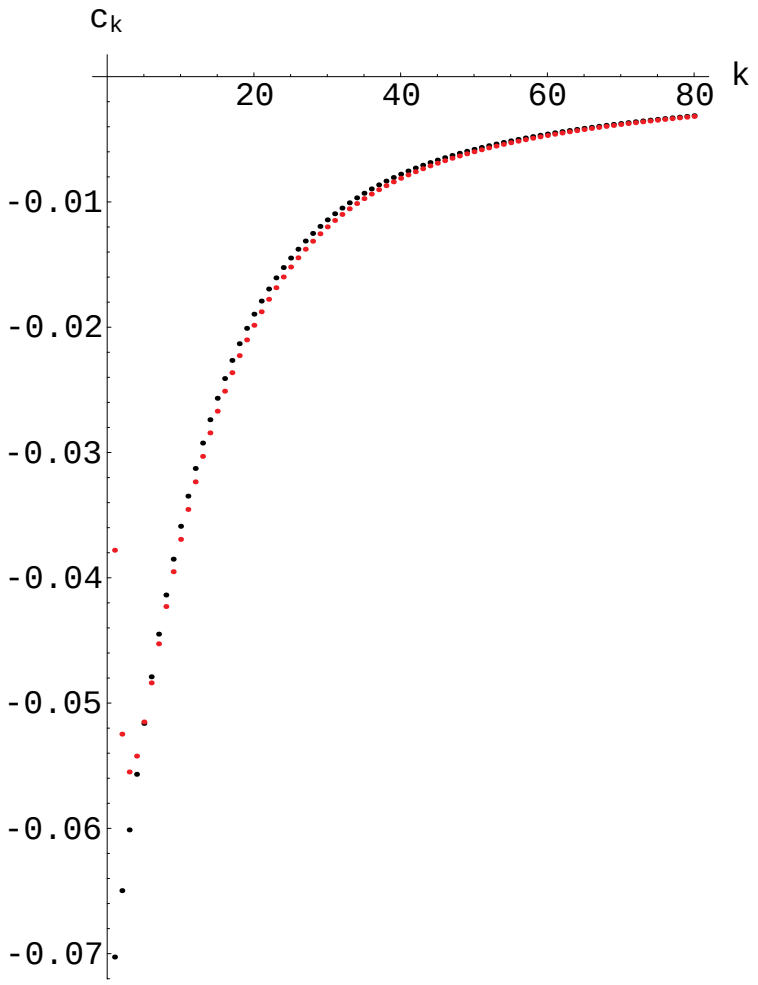


FIGURE 5. Plot of c_k for $\alpha = \frac{7}{2}, \beta = 4$ (black) vs. the Poisson approximation (red).

4.4. **The case $\alpha = \frac{1}{2}$.** In this case ($\alpha < 1$!) we cannot employ the argument of Appendix A, but we have for $\Re(s) - \epsilon \geq \rho = \frac{1}{2}$ and assuming $|c_k|$ increases with β:

$$\left|\varphi(s; \frac{1}{2}, \beta)\right| < \sum_{k=0}^{\infty} k^{-\left(\frac{\epsilon}{\beta}+1\right)} |c_k(1/2, \beta)| \leq \sum_{k=0}^{\infty} k^{-\left(\frac{\epsilon}{\beta}+1\right)} |c_k(1/2, \infty)|$$

From Subsection 4.2 we know that $|c_k(1/2,\infty)| = |\frac{1}{\zeta(1/2)} - 1| \cong 1.68477$, thus for any finite β, $\varphi(s;1/2,\beta)$ is also finite under the assumption that $\sup_\beta |c_k(1/2,\beta)|$ is bounded by $|c_k(1/2,\infty| \cong 1.68477$.

We remember that the great mathematician of the beginning of the century, J.F. Littlewood, has shown on the RH that the series

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n^{\frac{1}{2}+\epsilon}} \qquad \epsilon > 0$$

converges, even if with the Pochhammer's approach is not possible to have absolute convergence. We can verify the numerical bound of this series which should be smaller (or equal) than our predicted bound $A \cong 1.68477$ (Figure 6).

In the strong coupling limit we observe, with the help of our numerical results, the phenomena of a kind of "annihilation of the wave" in a macroscopic region of increasing width with β (α should be understood as $\alpha = \frac{1}{2}+$).

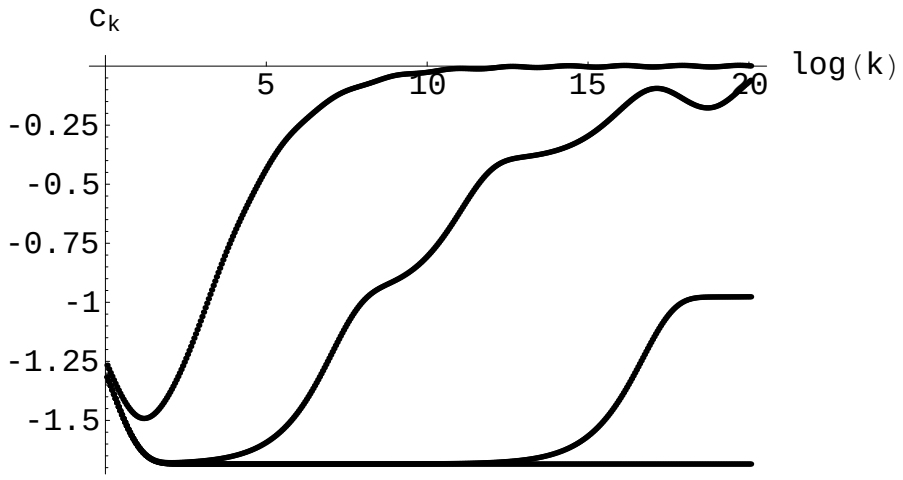


FIGURE 6. Plot of c_k for $\alpha = \frac{1}{2}, \beta = 4, 10, 24, 50$ from top to bottom (data obtained with $n = 10^6$ in (4)).

4.5. **Some cases with the same decay as the Hardy-Littlewood one.** We present some cases which should give the same behaviour as the original Hardy-Littlewood case ($\alpha = 1, \beta = 2$), then:

$$\alpha = \frac{\beta}{4} + \frac{1}{2}$$

For all the cases we present ($\beta = 6, 8, 10$), the c_k seem to decay as $k^{-\frac{1}{4}}$ up to $k = 500$ million (see Figure 7).

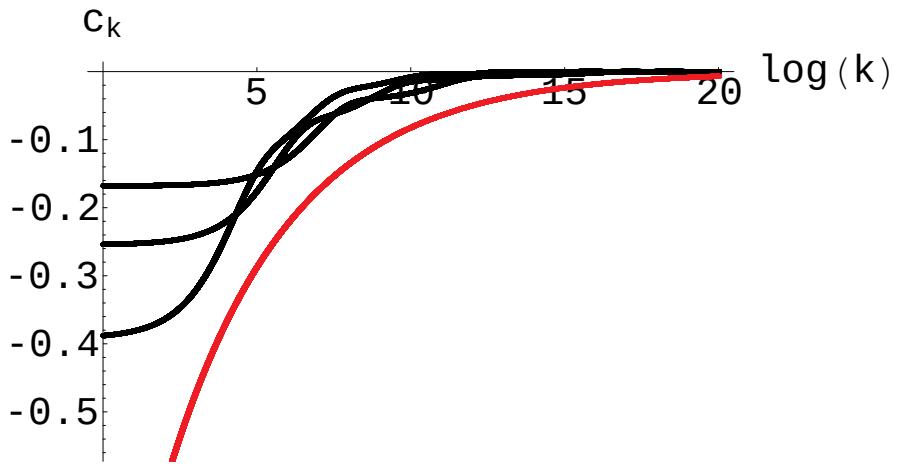


FIGURE 7. Plot of c_k if $\alpha = \frac{\beta}{4} + \frac{1}{2}$ for $\beta = 6, 8, 10$ [black curves from the bottom to the top] and the reference function $-k^{-\frac{1}{4}}$ [red curve] (data obtained with $n = 10^6$ in (4)).

5. Conclusions

In this work we have revisited the original criteria of Riesz and of Hardy-Littlewood for the Riemann Hypothesis in light of recent pioneering works concerning the possible representations of the Riemann Zeta function by means of the Pochhammer's polynomials. The discrete representation in the case $\alpha = \beta = 2$ is due to Baez-Duarte. In order to carry out our numerical experiments related to the criteria, we have first extended the analytical formulation to a more general class of sequences containing two parameters α and β; using a theorem of Baez-Duarte we have specified a sufficient and necessary condition for the truth of the RH for our general class of sequences i.e. for the decay of the coefficients c_k as a power law of k. Moreover in doing this we have found the emergence of a Poisson-like distribution for the c_k which should be exact in the large β limit. Numerical experiments have been carried out for various cases for low values of k.

(1) For $\alpha = \frac{7}{2}$ and $\beta = 4$ we have presented intensive calculation using the Möbius function up to $n = 10000$ and for k up to some hundreds. For this case, the power law decay $k^{-\frac{3}{4}}$ is the same as that appearing in the original work of Riesz ($\alpha = \beta = 2$) and also investigated numerically by Baez-Duarte. The experiments confirm the correctness of the power law within the range of the values of n and of k we were able to treat here. For large values of k see Footnote.
(2) For α and β such that the c_k should all give the power law decay $k^{-\frac{1}{4}}$ at large values of k to ensure the truth of the RH, i.e those where $\alpha = \frac{\beta}{4} + \frac{1}{2}$, we have presented experiments for some values of β which indicates this

power law decay. All sequences c_k have plots lying above a fixed curve of equation $y = Ak^{-\frac{1}{4}}$ for some fixed constant A independent of β, in the range of k we have considered.

(3) Finally we have considered some experiments in the large β limit which indicate that the plots of c_k become more and more flat, well approximated by the mean value of the Poisson-type distribution we have found. As β becomes large and large the c_k approaches in absolut value a constant, for all k, indicating that in this sense the RH may barely be true.

This work will be expanded with numerical experiments for bigger values of k [5] and in the search of other new representations of the Riemann Zeta function, different of the one considered here [6]. Moreover there is the aim that the new criteria will be useful in the context of additional numerical experiments. These works will be presented in a near future.

Appendix A

We follow strictly the lines of calculations of Baez-Duarte [2, 3] to show that the representation (6) for $[\zeta(s)]^{-1}$ is unconditionally valid for $\Re(s) > \rho = 1, \alpha > 1$ and $\beta > 0$. We consider the quantity:

$$q_k = \sum_{n=1}^{\infty} \frac{1}{n^{\alpha}} \left(1 - \frac{1}{n^{\beta}}\right)^k$$

Using the Euler-MacLaurin series (restricting ourselves to the main contribution), we have that:

$$q_k \cong \int_1^{\infty} \frac{1}{x^{\alpha}} \left(1 - \frac{1}{x^{\beta}}\right)^k dx$$

Then with the variable change $y = \frac{1}{x^{\beta}}$ we obtain:

$$\begin{aligned} q_k &\cong \frac{1}{\beta} \int_0^1 y^{\frac{\alpha-1}{\beta}-1} (1-y)^{k+1-1} dy \\ &= B(\frac{\alpha-1}{\beta}, k+1) \end{aligned}$$

where

$$B(\lambda, \mu) = \int_0^1 x^{\lambda-1} (1-x)^{\mu-1} dx = \frac{\Gamma(\lambda)\Gamma(\mu)}{\Gamma(\lambda+\mu)}$$

is the Beta function.

Thus for k large, we have

$$q_k \cong \frac{1}{\beta}\Gamma(\frac{\alpha-1}{\beta}) C \frac{k}{k^{\frac{\alpha-1}{\beta}+1}} \cong \frac{1}{k^{\frac{\alpha-1}{\beta}}}$$

Appendix B

Still following Baez-Duarte [2, 3] and here for the family of sequences with parameters α and β, we now show the necessity of the condition (7), assuming the RH to be true in the seminfinite strip $\Re(s) > \rho = \frac{1}{2}$.

We set $M(x) = \sum_{n \leq x} \mu(n)$, then we obtain $\forall \epsilon > 0$:

$$M(x) \leq x^{\rho+\epsilon}$$

Summation by parts gives for the main contribution:

$$|c_k| = \left| \int_1^\infty M(x) \frac{d}{dx} \left(\frac{1}{x^\alpha} \left(1 - \frac{1}{x^\beta} \right)^k \right) dx \right|$$

With the variable change $y = \frac{1}{x}$, using $M(\frac{1}{y}) \ll y^{-\rho-\epsilon}$ for $y \downarrow 0$ (RH) we have:

$$|c_k| \ll \alpha \int_0^1 y^{\alpha-\rho-\epsilon-1} (1-y^\beta)^k dy + \beta k \int_0^1 y^{\alpha+\beta-\rho-\epsilon-1} (1-y^\beta)^{k-1} dy$$

and finally with $y^\beta = z$ we obtain

$$|c_k| \ll \frac{\alpha}{\beta} \int_0^1 z^{\frac{\alpha-\rho-\epsilon}{\beta}-1} (1-z)^{k+1-1} dz + k \int_0^1 z^{\frac{\alpha-\rho-\epsilon+\beta}{\beta}-1} (1-z)^{k-1} dz$$

which for large k is given by:

$$|c_k| \ll \frac{\alpha}{\beta} \frac{\Gamma(\frac{\alpha-\rho-\epsilon}{\beta})}{k^{\frac{\alpha-\rho-\epsilon}{\beta}}} + \frac{\Gamma(\frac{\alpha-\rho-\epsilon+\beta}{\beta})}{k^{\frac{\alpha-\rho-\epsilon+\beta}{\beta}+1}} < \frac{C}{k^{\frac{\alpha-\rho-\epsilon}{\beta}}}$$

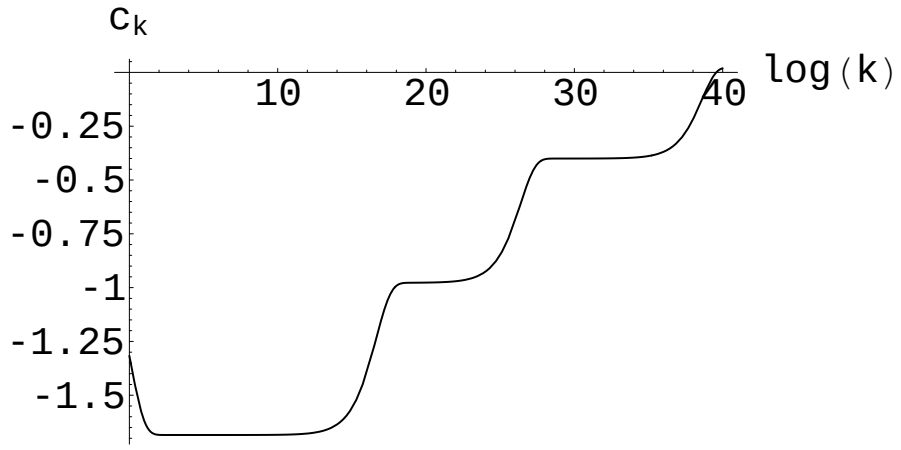


FIGURE 8. Plot of c_k for $\alpha = \frac{1}{2}, \beta = 24$ (data obtained with $n = 10^6$ in (4)).

FOOTNOTE

Our calculations have been carried out only to values of k not exceeding $k = 1000$ except for the Subsections 4.4 and 4.5. It is our pleasure to thank Prof. Luis Baez-Duarte for sending us, after the first draft of this paper, a copy of two recent published works by the author (in particular [3]), now added to our references. The paper contains the plot of the results of advanced numerical experiments up to $k = 100000$ by Krzystof Maslanka for the Riesz case, which clearly indicate that c_k becomes of oscillatory type with a wavelength related in first approximation to the first zero of the Riemann Zeta function.

We also thank Prof. Luis Baez-Duarte for sending us a picture of more numerical results, also concerning the Riesz case, by Marek Wolf, where values of k extend up to $k = 200000$ and confirming the oscillatory character of the sequence c_k, as well. Now a refinement of these results is published in arXiv [11].

We are currently performing an extension of our calculations, using the Poisson-Möbius formula considered in this work, for various cases, with the aim of obtaining satisfactory numerical results up to some billions for k and these will be presented in a forthcoming note. For the case $\alpha = \frac{1}{2}+, \beta = 24$, the plot of c_k up to $k = e^{40} \cong 2.3 \cdot 10^{17}$ is already given in Fig. 8 without comments, see only the emergence of plateau up to $\log(k) = 40$.

References

[1] Baez-Duarte L 2003 *arXiv:math.NT/0307214v1* 16 July 2003
[2] Baez-Duarte L 2003 *arXiv:math.NT/0307215v1* 16 July 2003
[3] Baez-Duarte L 2005 *International Journal of Mathematics and Mathematical Sciences* 2005:21 3527-3537
[4] Baez-Duarte L 2005 *arXiv:math.NT/0504402v1* 20 April 2005
[5] Beltraminelli S and Merlini D 2007 *in preparation, to be posted*
[6] Beltraminelli S and Merlini D 2007 *in preparation, to be posted*
[7] D'Errico M 2004 (unpublished) *presented at the International Workshop on Complex Systems (Cerfim-Issi)* held in Locarno (Switzerland), 16-18 september 2004
[8] Hardy GH and Littlewood JE 1918 *Acta Mathematica* **41** 119
[9] Maslanka K 2001 *arXiv:math-ph/0105007v1* 4 May 2001
[10] Riesz F 1916 *Acta Mathematica* **40** 185
[11] Wolf M 2006 *arXiv:math.NT/06054857v1* 17 May 2006

CERFIM, Research Center for Mathematics and Physics, PO Box 1132, 6600 Locarno, Switzerland
E-mail address: stefano.beltraminelli@ti.ch
E-mail address: merlini@cerfim.ch

ALBANIAN JOURNAL
OF MATHEMATICS
Volume 1, Number 1, Pages 31–56
ISSN 1930-1235(electronic version)

L_p-NORM GENERALISED SYMMETRISED DIRICHLET DISTRIBUTIONS

ENKELEJD HASHORVA, SAMUEL KOTZ, AND ALFRED KUME

ABSTRACT. The paper deals with random vectors $\boldsymbol{X}$ possessing the stochastic representation $\boldsymbol{X} = R\boldsymbol{\mathcal{U}}$ where R is a positive random radius and $\boldsymbol{\mathcal{U}}$ is a L_p-norm generalised symmetrised Dirichlet random vector independent of R. The Kotz Type I multivariate distribution appears prominently in the asymptotic results.

1. INTRODUCTION

Let $\boldsymbol{X}$ be a spherical random vector in $\mathbb{R}^k, k \geq 2$, i.e. the distribution function of $\boldsymbol{X}$ is invariant with respect to orthogonal transformations in $\mathbb{R}^k$. Define the associated random radius R by the stochastic representation $R \stackrel{d}{=} (\sum_{i=1}^k X_i^2)^{1/2}$ ($\stackrel{d}{=}$ stands for equality in distribution). Cambanis et al. (1981) show in their pioneering paper that if $R > 0$ almost surely we have the stochastic representation

$$\boldsymbol{X} \stackrel{d}{=} R\boldsymbol{\mathcal{U}}, \tag{1}$$

with $\boldsymbol{\mathcal{U}}$ uniformly distributed on the unit sphere of $\mathbb{R}^k$ independent of the associated random radius R.

The main distributional properties of elliptical random vectors can be found in Kotz (1975), Cambanis et al. (1981), Anderson and Fang (1990), Fang et. al (1990), Fang and Zhang (1990), Szabłowski (1990), Berman (1992), Gupta and Varga (1993), Kano (1994), Kotz and Ostrovskii (1994) among many other sources.

When $\boldsymbol{\mathcal{U}}$ is uniformly distributed on the unit sphere of $\mathbb{R}^k$ the spherical distributions become quite tractable. On the other hand, due to these restrictions some important multivariate distributions such as Dirichlet distributions with unequal parameters do not belong to this class. Eliminating the assumption of the uniformity of the distribution function of $\boldsymbol{\mathcal{U}}$ allows studying more general distributions which share the simple stochastic representation (1).

Fang and Fang (1990) chose $\boldsymbol{\mathcal{U}}$ to have generalised symmetrised Dirichlet distribution (see below (2) for the definition) thus introducing generalised symmetrised Dirichlet random vectors in $\mathbb{R}^k$ with the stochastic representation (1). In the aforementioned paper several properties of this new class of random vectors are given. The well known in the theory and practice Dirichlet distribution was originally introduced by P.G.L. Dirichlet (a famous French-German mathematician in 1839).

Received by the editors September 11, 2006 and, in revised form, February 10, 2007.
2000 *Mathematics Subject Classification.* Primary 60F05; Secondary 60G70.
Key words and phrases. L_p-norm generalised symmetrised Dirichlet distributions, tail asymptotics, conditional limiting theorems, max-domain of attractions.

Another possible generalisation is to deal with the general L_p-norm ($p > 0$), but still retain the condition that $\boldsymbol{\mathcal{U}}$ is uniformly distributed on the unit sphere of $\mathbb{R}^k$ with respect to the L_p-norm. This approach is suggested by Gupta and Song (1997) and Szabłowski (1998). Distributional properties of L_p-norm spherical random vectors derived in Gupta and Song (1997) and Szabłowski (1998) are, as expected, similar to the properties derived for L_2-norm spherical random vectors in Cambanis et al. (1981). Fang and Fang (1990), Gupta and Song (1997) and Szabłowski (1998) provide results that are shared by a wide class of multivariate distribution functions, and in particular by the class of spherical random vectors.

In this paper we shall consider a further generalisation (combining results in the aforementioned papers) by taking $\boldsymbol{\mathcal{U}}$ in (1) to be a multivariate L_p-norm generalised symmetrised Dirichlet distribution (see below (2)), introducing thus via (1) L_p-norm generalised symmetrised Dirichlet random vectors.

We provide in this paper some basic distributional properties of L_p-norm generalised symmetrised Dirichlet (LpGSD) distributions. Furthermore, we obtain certain asymptotic results which are in line with the previous results for L_p-norm spherical random vectors. Conditional limiting theorems are derived in the last section. It is quite surprising that the standard Kotz Type I LpGSD distribution approximates a large subclass of LpGSD random vectors.

2. Notation and Preliminaries

For completeness we shall first present some notation and then review several known results about L_p-norm spherical random vectors and generalised symmetrised Dirichlet ones.

Let I be a non-empty subset of $\{1,\ldots,k\}, k \geq 2$, and set $J := \{1,\ldots,k\} \setminus I$. For any vector $\boldsymbol{x} = (x_1,\ldots,x_k)^\top \in \mathbb{R}^k$ set $\boldsymbol{x}_I := (x_i, i \in I)^\top$, and write $\boldsymbol{x}_I^\top$ in place of $(\boldsymbol{x}_I)^\top$. Denote for two vectors in $\mathbb{R}^k, \boldsymbol{x}, \boldsymbol{y}$ the operations

$$\begin{aligned}
&\boldsymbol{x}+\boldsymbol{y} := (x_1+y_1,\ldots,x_k+y_k),\\
&\boldsymbol{x} > \boldsymbol{y}, \text{ if } x_i > y_i, \quad \forall\, i=1,\ldots,k,\\
&\boldsymbol{x} \geq \boldsymbol{y}, \text{ if } x_i \geq y_i, \quad \forall\, i=1,\ldots,k,\\
&\boldsymbol{x} \neq \boldsymbol{y}, \text{ if for some } i \leq k,\, x_i \neq y_i,\\
&\boldsymbol{x} \not\leq \boldsymbol{y}, \text{ if for some } i \leq k, x_i > y_i,
\end{aligned}$$

and define

$$\begin{aligned}
&\boldsymbol{a}\boldsymbol{x} := (a_1x_1,\ldots,a_kx_k)^\top, \quad c\boldsymbol{x} := (cx_1,\ldots,cx_k)^\top, \quad \boldsymbol{a} \in \mathbb{R}^k, c \in \mathbb{R},\\
&\|\boldsymbol{x}_I\|_p := \Big(\sum_{i\in I}|x_i|^p\Big)^{1/p}, \quad p>0, \quad (L_p\text{-norm}),\\
&\mathcal{S}_p^{k-1} := \{\boldsymbol{x} \in \mathbb{R}^k : \|\boldsymbol{x}\|_p = 1\}, \quad (\text{unit sphere}).
\end{aligned}$$

Let moreover $\mathbf{0} := (0,\ldots,0)^\top \in \mathbb{R}^k$, $\mathbf{1} := (1,\ldots,1)^\top \in \mathbb{R}^k$. We shall be denoting by $Beta(a,b)$ and $Gamma(a,b)$ respectively, the distribution functions of a Beta or a Gamma random variables with parameters a and b. If a random vector $\boldsymbol{Z}$ has the distribution function Q, this will be indicated by $\boldsymbol{Z} \sim Q$.

Throughout the paper $\boldsymbol{\alpha} := (\alpha_1, \ldots, \alpha_k), k \geq 2$ will denote a vector with positive components i.e. $\boldsymbol{\alpha} > 0$, p be a fixed positive constant ($p > 0$) and

$$\overline{\alpha} := \sum_{i=1}^{k} \alpha_i, \quad \overline{\alpha}_I := \sum_{i \in I} \alpha_i, \quad I \subset \{1, \ldots, k\}, |I| \geq 1.$$

The probability density function (p.d.f) of the Dirichlet distribution (see e.g. Kotz et al. (2000)) is given by

$$\frac{\Gamma(\overline{\alpha})}{\prod_{i=1}^{k} \Gamma(\alpha_i)} \Big(1 - \sum_{i=1}^{k-1} u_i\Big)^{a_k - 1} \prod_{i=1}^{k-1} u_i^{\alpha_i - 1},$$

where $\sum_{i=1}^{k-1} u_i \leq 1, u_i > 0, i = 1, \ldots, k$. Denote by $(U_1, \ldots, U_{k-1})^\top$ a random vector with the above p.d.f and write $(U_1, \ldots, U_{k-1})^\top \sim \mathcal{D}(k, \boldsymbol{\alpha})$. The transformed random vector $(U_1^{1/p}, \ldots, U_{k-1}^{1/p})^\top$ has p.d.f

$$\frac{p^{k-1}\Gamma(\overline{\alpha})}{\prod_{i=1}^{k} \Gamma(\alpha_i)} \Big(1 - \sum_{i=1}^{k-1} u_i^p\Big)^{a_k - 1} \prod_{i=1}^{k-1} u_i^{p\alpha_i - 1}, \quad u_i > 0, i = 1, \ldots, k : \sum_{i=1}^{k-1} |u_i|^p \leq 1.$$

Let $\mathcal{I}_1, \ldots, \mathcal{I}_k$ be independent random variables taking values $-1, 1$ with probability $1/2$. The random vector $(\mathcal{I}_1 U_1^{1/p}, , \ldots, \mathcal{I}_{k-1} U_{k-1}^{1/p})^\top$ represents a symmetrisation with power p of $(U_1, \ldots, U_{k-1})^\top$. (For $p = 2$ it is referred in literature simply as symmetrisation). The p.d.f of the symmetrised random vector is thus

$$h(u_1, \ldots, u_{k-1}) := \frac{p^{k-1}\Gamma(\overline{\alpha})}{2^{k-1} \prod_{i=1}^{k} \Gamma(\alpha_i)} \Big(1 - \sum_{i=1}^{k-1} |u_i|^p\Big)^{a_k - 1} \prod_{i=1}^{k-1} |u_i|^{p\alpha_i - 1}, \tag{2}$$

where $\sum_{i=1}^{k-1} |u_i|^p \leq 1$. (See (8) below for alternative derivation of this p.d.f).
Fang and Fang (1990) designate $\boldsymbol{\mathcal{U}} = (\mathcal{U}_1, \ldots, \mathcal{U}_k)^\top$ to have symmetrised Dirichlet distribution (with respect to $\boldsymbol{\alpha}$) provided $\|\boldsymbol{\mathcal{U}}\|_2 = 1$ and $\mathcal{U}_i = \mathcal{I}_i U_i^{1/2}, i = 1, \ldots, k - 1$. We shall extend that definition as follows.

Definition 1. A random vector $\boldsymbol{\mathcal{U}}$ in $\mathbb{R}^k$ is said to have the L_p-norm symmetrised Dirichlet (LpSD) distribution with parameter $\boldsymbol{\alpha}$ if $\|\boldsymbol{\mathcal{U}}\|_p = 1$ almost surely and $(\mathcal{U}_1, \ldots, \mathcal{U}_{k-1})^\top$ has the p.d.f h given by (2). (We shall denote $\boldsymbol{\mathcal{U}} \sim \mathcal{SD}(k, p, \boldsymbol{\alpha})$).

In some cases it may be more convenient to utilise unsymmetrised Dirichlet distributions. We denote

$$\boldsymbol{\mathcal{U}} \sim \mathcal{D}(k, p, \boldsymbol{\alpha}) \tag{3}$$

if $\boldsymbol{\mathcal{U}} \geq \mathbf{0}$ and $\|\boldsymbol{\mathcal{U}}\|_p = 1$ almost surely such that $(\mathcal{U}_1, \ldots, \mathcal{U}_{k-1})^\top$ has the density function $2^{k-1} h(u_1, \ldots, u_{k-1}), u_i > 0, i \leq k$, where h is defined in (2).

Considering $\boldsymbol{\mathcal{U}}$ to be a L_p-norm symmetrised Dirichlet random vector with the stochastic representation (1) we arrive at the following definition.

Definition 2. A random vector $\boldsymbol{X}$ in $\mathbb{R}^k, k \geq 2$ is said to possess a L_p-norm generalised symmetrised Dirichlet distribution with parameter $\boldsymbol{\alpha}$ (denoted $\boldsymbol{X} \sim \mathcal{GSD}(k, p, \boldsymbol{\alpha}, F)$) if it possesses stochastic representation (1) where $R > 0$, almost surely with the distribution function F independent of $\boldsymbol{\mathcal{U}}$, where $\boldsymbol{\mathcal{U}} \sim \mathcal{SD}(k, p, \boldsymbol{\alpha})$.

In the next section we shall derive some basic properties of the LpGSD random vectors and then proceed to Section 4 where we shall discuss dependence and asymptotic dependence of LpGSD distributions. Conditional limiting theorems motivated by previous results in Berman (1982,1983) and Berman (1992) are derived in Section 5. Section 5 focuses on the asymptotic tail behaviour in the case when the associated random radius is regularly varying. The proofs are relegated to Section 7. Further theoretical results are provided in the Appendix.

3. Main Distributional Properties

Using the derivations and definitions presented in Section 2 for a random vector $\boldsymbol{X} \sim \mathcal{GSD}(k, p, \boldsymbol{\alpha}, F)$ we have the following stochastic representation

$$\boldsymbol{X} \stackrel{d}{=} R\boldsymbol{\mathcal{U}} \stackrel{d}{=} R(\mathcal{I}_1 U_1^{1/p}, \ldots, \mathcal{I}_{k-1} U_{k-1}^{1/p}, \mathcal{I}_k \mathcal{U}_k)^\top, \tag{4}$$

where $(U_1, \ldots, U_{k-1})^\top$ is a Dirichlet random vector with parameter $\boldsymbol{\alpha}$ and $\mathcal{U}_k > 0$ is such that the relation

$$\sum_{i=1}^{k-1} U_i + \mathcal{U}_k^p \quad = \quad 1$$

is valid almost surely.

The stochastic representation (3.1) shows that the role of parameter p in the distributional properties of LpGSD random vectors is determined solely by the power transformation of the Dirichlet random vector $(U_1, \ldots, U_{k-1})^\top$. Although the distributional properties of Dirichlet random vectors are well-known (see e.g. Fang et al. (1990) or Kotz et al. (2000)), the main properties of LpGSD random vectors do not follow automatically. Further derivations are needed (as in Fang and Fang (1990)) to obtain the main distributional properties. Gupta and Song (1997) have shown that the L_p-norm spherical random vectors possess the same properties as the L_2-norm spherical (or simply spherical) random vectors. Here we shall show that the same is valid for LpGSD random vectors, utilising the techniques presented by Fang and Fang (1990).

First we shall observe that it is possible to arrive at the definition of the LpGSD random vectors via a single density generator (as it is presented in Fang and Fang (1990) for the L_2-norm).

Actually the definition of the density generator is unrelated to p and is therefore similar to the one given in Fang and Fang (1990) presented below:

Definition 3. (Density generator) Let g be a positive measurable function, and $\boldsymbol{\alpha} := (\alpha_1, \ldots, \alpha_k)^\top, k \geq 2$ be a given vector with positive components. If for some $\omega \in (0, \infty]$ the function g satisfies

$$\left(\frac{2}{p}\right)^k \frac{\prod_{i=1}^k \Gamma(\alpha_i)}{\Gamma(\overline{\alpha})} \int_0^\omega g(x) x^{\overline{\alpha}-1}\, dx = 1, \quad \overline{\alpha} := \sum_{i=1}^k \alpha_i, \tag{5}$$

we shall call g to be a density generator with respect to $\boldsymbol{\alpha}$, denoting $g \sim \mathcal{G}(\boldsymbol{\alpha}, \omega)$. If the integral above is finite for any $\boldsymbol{\alpha}$ (with positive components) we shall refer to g as the universal density generator.

The next result shows that a density generator uniquely defines the p.d.f of a LpGSD random vector.

Theorem 1. *Let $g \sim \mathcal{G}(\boldsymbol{\alpha}, \omega)$ with $\boldsymbol{\alpha} \in (0,\infty)^k, k \geq 2$, and $\omega \in (0,\infty]$, and $\boldsymbol{X}$ be a k-dimensional random vector with the density function h defined by*

$$h(\boldsymbol{x}) := g(\sum_{i=1}^{k} |x_i|^p) \prod_{i=1}^{k} |x_i|^{p\alpha_i - 1}, \quad \forall \boldsymbol{x} \in \mathbb{R}^k : 0 < \|\boldsymbol{x}\|_p < \omega, \tag{6}$$

with $p > 0$. Then $\boldsymbol{X} \sim \mathcal{GSD}(k, p, \boldsymbol{\alpha}, F)$ where F is a distribution function on $[0, \omega)$ with the p.d.f f

$$f(r) \;=\; 2\left(\frac{2}{p}\right)^{k-1} \frac{\prod_{i=1}^{k} \Gamma(\alpha_i)}{\Gamma(\overline{\alpha})} g(r^p) r^{p\overline{\alpha}-1}, \quad \forall r \in (0, \omega). \tag{7}$$

Conversely, if $\boldsymbol{X} \sim \mathcal{GSD}(k, p, \boldsymbol{\alpha}, F)$ with F being a distribution function with the p.d.f f then $\boldsymbol{X}$ possesses the density function h defined in (6) *with the density generator g defined by the density f in* (7)*.*

In the case when $\boldsymbol{X}$ is defined in terms of a density generator $g \sim \mathcal{G}(\boldsymbol{\alpha}, \omega)$ we shall denote $\boldsymbol{X} \sim \mathcal{GSD}(k, p, \boldsymbol{\alpha}, g)$ suppressing the symbol ω.

Several examples below should clarify the definitions and the theorem above.

Example 1. [**Symmetrised Dirichlet**] Let $\boldsymbol{\alpha} \in (0,\infty)^k, k \geq 2$ and c, p be positive constants and specify $g(x) = c(1-x)^{\alpha_k - 1}, \forall x \in (0,1)$. Define the density function h of a random vector $(\mathcal{U}_1, \ldots, \mathcal{U}_{k-1})^\top$ in $\mathbb{R}^{k-1}$ as in (6) by

$$h(\boldsymbol{x}) := c(1 - \sum_{i=1}^{k-1} |x_i|^p)^{\alpha_k - 1} \prod_{i=1}^{k-1} |x_i|^{p\alpha_i - 1}, \quad \boldsymbol{x} \in \mathbb{R}^{k-1} : 0 < \|\boldsymbol{x}\|_p^p < 1.$$

Utilising 5 we arrive at:

$$\begin{aligned} c^{-1} \;&=\; \left(\frac{2}{p}\right)^{k-1} \frac{\prod_{i=1}^{k-1} \Gamma(\alpha_i)}{\Gamma(\sum_{i=1}^{k-1} \alpha_i)} \int_0^1 x^{\sum_{i=1}^{k-1} \alpha_i - 1} (1-x)^{\alpha_k - 1}\, dx \\ &=\; \left(\frac{2}{p}\right)^{k-1} \frac{\prod_{i=1}^{k} \Gamma(\alpha_i)}{\Gamma(\overline{\alpha})}, \quad \overline{\alpha} := \sum_{i=1}^{k} \alpha_i. \end{aligned}$$

Consequently we have for any $\boldsymbol{x} \in \mathbb{R}^{k-1}$ such that $\|\boldsymbol{x}\|_p < 1$

$$h(x_1, \ldots, x_{k-1}) \;=\; \left(\frac{p}{2}\right)^{k-1} \frac{\Gamma(\overline{\alpha})}{\prod_{i=1}^{k} \Gamma(\alpha_i)} (1 - \sum_{i=1}^{k-1} |x_i|^p)^{\alpha_k - 1} \prod_{i=1}^{k-1} |x_i|^{p\alpha_i - 1}. \tag{8}$$

Example 2. [**Kotz Type I**] Let the density generator g be of the form

$$g(x) \;=\; c x^N \exp(-r x^s), \quad x > 0, c > 0, N \in \mathbb{R}, r > 0, s > 0. \tag{9}$$

For a given $\boldsymbol{\alpha}$ and restricting $N > -\overline{\alpha}$ we obtain using (5) that the constant c is determined by

$$c\left(\frac{2}{p}\right)^{k} \frac{\prod_{i=1}^{k} \Gamma(\alpha_i)}{\Gamma(\overline{\alpha})} \int_0^\infty x^N \exp(-r x^s) x^{\overline{\alpha}-1}\, dx \;=\; 1.$$

Observing that

$$\int_0^\infty x^N \exp(-r x^s) x^{\overline{\alpha}-1}\, dx \;=\; \frac{\Gamma((N+\overline{\alpha})/s)}{s r^{(N+\overline{\alpha})/s}} \tag{10}$$

we arrive at:

$$c := \left(\frac{p}{2}\right)^k \frac{sr^{(N+\overline{\alpha})/s}}{\Gamma((N+\overline{\alpha})/s)} \frac{\Gamma(\overline{\alpha})}{\prod_{i=1}^k \Gamma(\alpha_i)}. \tag{11}$$

Here the density generator g given by (9) is a universal one.
We say that $\boldsymbol{X}$ in $\mathbb{R}^k$ is a Kotz Type I LpGSD random vector if its density function h is given for any $\boldsymbol{x} \in \mathbb{R}^k$ by

$$h(\boldsymbol{x}) := \left(\frac{p}{2}\right)^k \frac{sr^{(N+\overline{\alpha})/s}}{\Gamma((N+\overline{\alpha})/s)} \frac{\Gamma(\overline{\alpha})}{\prod_{i=1}^k \Gamma(\alpha_i)} \|\boldsymbol{x}\|_p^{pN} \exp(-r\|\boldsymbol{x}\|_p^{ps}) \prod_{i=1}^k |x_i|^{p\alpha_i - 1}, \tag{12}$$

involving the norm of $\boldsymbol{x}$, an exponential function and a product of the components of $\boldsymbol{x}$ (compare with (8)). In the standardised case $N = 0$ and $r = s = 1$ the random vector $\boldsymbol{X}$ possesses independent components such that

$$|X_i|^p \sim Gamma(\alpha_i, 1/p), \quad \forall i = 1, \ldots, k. \tag{13}$$

We shall denote by $\boldsymbol{\mathcal{K}}_{\boldsymbol{\alpha},p}$ the distribution function of $\boldsymbol{X}$ when $N + 1 = r = s = 1$.

Example 3. [**Kotz Type II**] Let the density generator g be of the form

$$g(x) = cx^N \exp(-rx^s), \quad x > 0, c > 0, N < 0, r > 0, s < 0. \tag{14}$$

Here the values of N and s are negative. Analogously to the previous example, for a given $\boldsymbol{\alpha}$, the constant c is obtained from (5)

$$c\left(\frac{2}{p}\right)^k \frac{\prod_{i=1}^k \Gamma(\alpha_i)}{\Gamma(\overline{\alpha})} \int_0^\infty x^N \exp(-rx^s) x^{\overline{\alpha}-1}\, dx = 1.$$

Choosing $N < -\overline{\alpha}$ we obtain

$$c := \left(\frac{p}{2}\right)^k \frac{(-s)r^{(N+\overline{\alpha})/s}}{\Gamma((N+\overline{\alpha})/s)} \frac{\Gamma(\overline{\alpha})}{\prod_{i=1}^k \Gamma(\alpha_i)} > 0.$$

Here g is also a universal density generator.
We define $\boldsymbol{X}$ in $\mathbb{R}^k$ to be a Kotz Type II LpGSD random vector provided its p.d.f h is given for $N < -\overline{\alpha}$ and $\boldsymbol{x} \in \mathbb{R}^k$ by

$$h(\boldsymbol{x}) := \left(\frac{p}{2}\right)^k \frac{|s|r^{(N+\overline{\alpha})/s}}{\Gamma((N+\overline{\alpha})/s)} \frac{\Gamma(\overline{\alpha})}{\prod_{i=1}^k \Gamma(\alpha_i)} \|\boldsymbol{x}\|_p^{pN} \exp(-r\|\boldsymbol{x}\|_p^{ps}) \prod_{i=1}^k |x_i|^{p\alpha_i - 1}.$$

The L_p-norm Kotz Type II spherical random vectors are considered in Hashorva (2006d). The original definition of these random vectors for the L_2-norm case is due to Kotz (1975).

Example 4. [**Kotz Type III**] Let $\boldsymbol{X} = R\boldsymbol{U}$ with R a positive random radius independent of the k-dimensional random vector $\boldsymbol{U}$ which is such that $\|\boldsymbol{U}\|_p = 1$ almost surely. We refer to $\boldsymbol{X}$ as a Kotz Type III random vector if the associated random radius $R > 0$ has asymptotic tail behaviour ($u \to \infty$)

$$\boldsymbol{P}\{R > u\} = (1 + o(1))Ku^N \exp(-ru^\delta) \quad K > 0, \delta \in \mathbb{R}, N \in \mathbb{R}, r > 0. \tag{15}$$

For $\delta \leq 0$ we assume that $N < 0$. If $\boldsymbol{U}$ is a LpGSD random vector then $\boldsymbol{X}$ is a LpGSD random vector.
Both Kotz Type I and Type II LpGSD random vectors belong to the larger class of the Kotz Type III random vectors.

Example 5. [**Pearson Type VII**] The density generator is $g(x) = c(1 + t/s)^{-N}$ with c, s positive constants. Assuming $N > \overline{\alpha}$ we obtain the density function h of a k-dimensional LpGSD Pearson Type VII distribution

$$h(\boldsymbol{x}) \quad = \quad \left(\frac{p}{2}\right)^k s^{-\overline{\alpha}} \frac{\Gamma(N)}{\Gamma(N-\overline{\alpha})\prod_{i=1}^k \Gamma(\alpha_i)} (1 + \sum_{i=1}^k |x_i|^p/s)^N \prod_{i=1}^k |x_i|^{p\alpha_i - 1},$$

for all $\boldsymbol{x} \in \mathbb{R}^k$.

Example 6. [**Kummer-Beta**] Let g be a density generator of a Kummer-Beta LpGSD distribution given by

$$g(x) \quad = \quad cx^{\delta-1}(1-x)^{\gamma-1}\exp(-\lambda x), \quad 0 < x < 1, \delta > 0, \lambda \geq 0, \gamma > 0.$$

The normalising constant c for given positive constants $\alpha_i, i \leq k$ such that $\overline{\alpha} > 1-\delta$ is specifically determined via the relations:

$$\begin{aligned} c^{-1} &= \left(\frac{2}{p}\right)^k \frac{\prod_{i=1}^k \Gamma(\alpha_i)}{\Gamma(\overline{\alpha})} \int_0^1 x^{\overline{\alpha}-1}\exp(-\lambda x) x^{\delta-1}(1-x)^{\gamma-1}\, dx \\ &= \left(\frac{2}{p}\right)^k \frac{\prod_{i=1}^k \Gamma(\alpha_i)}{\Gamma(\overline{\alpha})} \int_0^1 x^{\overline{\alpha}+\delta-2}(1-x)^{\gamma-1}\exp(-\lambda x)\, dx \\ &= \left(\frac{2}{p}\right)^k \frac{\prod_{i=1}^k \Gamma(\alpha_i)}{\Gamma(\overline{\alpha})} \frac{{}_1F_1(\overline{\alpha}+\delta-1; \overline{\alpha}+\delta+\gamma-1; -\lambda)\Gamma(\overline{\alpha}+\delta-1)\Gamma(\gamma)}{\Gamma(\overline{\alpha}+\delta+\gamma-1)}, \end{aligned}$$

where ${}_1F_1$ is the confluent hypergeometric function of the first kind (also known as Kummer's function of the first kind). ${}_1F_1$ has a hypergeometric series expansion given by

$${}_1F_1(a,b,x) = 1 + \frac{a}{b}x + \frac{a(a+1)}{b(b+1)}\frac{x^2}{2!} + \cdots = \sum_{k=0}^{\infty} \frac{(a)_k}{(b)_k}\frac{x^k}{k!},$$

where $(a)_k, (b)_k$ are the Pochhammer symbols.

Example 7. [**Kummer-Gamma**] The density generator g of a Kummer-Gamma LpGSD distribution is specified as

$$g(x) \quad = \quad cx^{\delta-1}(1+x)^{\gamma-1}\exp(-\lambda x), \quad x > 0, \delta > 0, \lambda > 0, \gamma > 0,$$

with

$$\begin{aligned} c^{-1} &= \left(\frac{2}{p}\right)^k \frac{\prod_{i=1}^k \Gamma(\alpha_i)}{\Gamma(\overline{\alpha})} \int_0^\infty x^{\overline{\alpha}-1}\exp(-\lambda x) x^{\delta-1}(1+x)^{\gamma-1}\, dx \\ &= \left(\frac{2}{p}\right)^k \frac{\prod_{i=1}^k \Gamma(\alpha_i)}{\Gamma(\overline{\alpha})} \Psi(\overline{\alpha}+\delta-1; \overline{\alpha}+\delta-\gamma-1; \lambda)\Gamma(\overline{\alpha}+\delta-1), \end{aligned}$$

where Ψ is the confluent hypergeometric function of the second kind. It is also known as the Kummer's function of the second kind, Tricomi function, or Gordon function.

See Kotz and Ng (1995) for some basic properties of the Kummer-Beta and Kummer-Gamma distributions. We note in passing that a Kummer-Gamma LpGSD random vector belongs to the class of Kotz Type III LpGSD random vectors defined in Example 4.

In addition to the fulfillment of the stochastic representation (1) the most distinguishing property of LpGSD distributions is the so-called amalgamation property, initially presented in Cambanis at al. (1981) for elliptical random vectors, and in

Fang and Fang (1990), Gupta and Song (1997) for generalised symmetrised Dirichlet and L_p-norm spherical random vectors, respectively.

Theorem 2 (Amalgamation property)**.** *Let $I_1,\ldots,I_m, m\geq 2$ be a partition of $\{1,\ldots,k\}, k\geq 1$ and $\boldsymbol{X}\sim \mathcal{GSD}(k,p,\boldsymbol{\alpha},F)$ be a k-dimensional random vector as in Theorem 1. Then for any $j=1,\ldots,m$ we have the stochastic representation*

$$\boldsymbol{X}_{I_j} \stackrel{d}{=} R W_j \boldsymbol{Z}_{I_j}, \tag{16}$$

where the variables $R, \boldsymbol{W}_m:=(W_1,\ldots,W_m)^\top, \boldsymbol{Z}_{I_1},\ldots,\boldsymbol{Z}_{I_m}$ are pairwise independent random vectors with the random variable $R>0$, and

$$R\sim F, \quad \boldsymbol{W}_m\sim \mathcal{D}(m,p,\boldsymbol{a}_m), \quad \boldsymbol{Z}_{I_j}\sim \mathcal{SD}(k_j,p,\boldsymbol{\alpha}_{I_j}), \quad j=1,\ldots,m, \tag{17}$$

with $\boldsymbol{a}_m:=(\sum_{i\in I_1}\alpha_i,\ldots,\sum_{i\in I_m}\alpha_i)^\top, k_j=|I_j|\geq 1$ and $\mathcal{SD}(k_j,p,\boldsymbol{\alpha}_{I_j})$, $\mathcal{D}(m,p,\boldsymbol{a})$ as in Definition 2.1 and 2.2, respectively.

For any non-empty index set I and $p>0$ we define the associated random radius $R_{I,p}$ of $\boldsymbol{X}$ by

$$R_{I,p}:=\Big(\sum_{i\in I}|X_i|^p\Big)^{1/p}=\|\boldsymbol{X}_I\|_p>0, \quad p>0.$$

In the case $I=\{1,\ldots,k\}$ we shall simply write R instead of $R_{I,p}$.

Corollary 3. *Let $\boldsymbol{X}$ be defined as in Theorem 2, and $R_{I,p}$ be the associated random radius of $\boldsymbol{X}$ with respect to the non-empty index set I with m elements. Then the stochastic representation*

$$\boldsymbol{X}_I \stackrel{d}{=} R_{I,p}\boldsymbol{\mathcal{V}}_I \tag{18}$$

is valid with $R_{I,p}$ independent of $\boldsymbol{\mathcal{V}}_I$ where $\boldsymbol{\mathcal{V}}_I\sim\mathcal{SD}(m,p,\boldsymbol{\alpha}_I)$. Furthermore, if $m<k$

$$R_{I,p}^p \stackrel{d}{=} R^p W \tag{19}$$

holds where $W>0$ is distributed as $W\sim Beta(\overline{\alpha}_I,\overline{\alpha}-\overline{\alpha}_I)$, with W,R independent.

Corollary 4. *Let $\boldsymbol{X}\sim\mathcal{GSD}(k,p,\boldsymbol{\alpha},F)$ be a LpGSD random vector in $\mathbb{R}^k$. Then we have the stochastic representation*

$$X_j\stackrel{d}{=} R\mathcal{I}_j\Big[|\cos(\Theta_j)|\prod_{i=1}^{j-1}\sin(\Theta_i)\Big]^{2/p}, \quad 1\leq j\leq k-1, \tag{20}$$

$$X_k\stackrel{d}{=} R\mathcal{I}_k\Big[|\sin(\Theta_{k-1})|\prod_{i=1}^{k-2}\sin(\Theta_i)\Big]^{2/p}, \tag{21}$$

where $\mathcal{I}_j=\mathrm{sign}(\cos(\Theta_j)), 1\leq j\leq k-1, \mathcal{I}_k=\mathrm{sign}(\sin(\Theta_{k-1}))$ are independent random variables, being further independent of the random angles $\Theta_i, 1\leq i\leq k-1$ which have the density functions

$$q_i(\theta):=\frac{\Gamma(\overline{\alpha}_J)}{\Gamma(\overline{\alpha}_J-\alpha_i)\Gamma(\alpha_i)}|\sin(\theta)|^{2\overline{\alpha}_J-1}|\cos(\theta)|^{2\alpha_i-1}, \quad 0\leq\theta\leq\pi, 1\leq i\leq k-2,$$

where $J:=\overline{\alpha}-\sum_{j=1}^i\alpha_i$, and

$$q_{k-1}(\theta):=\frac{1}{2}\frac{\Gamma(\alpha_{k-1}+\alpha_k)}{\Gamma(\alpha_{k-1})\Gamma(\alpha_k)}|\sin(\theta)|^{2\alpha_k-1}|\cos(\theta)|^{2\alpha_{k-1}-1}, \quad 0\leq\theta\leq 2\pi.$$

Furthermore, $R, \Theta_i, 1 \le i \le k-1$ are independent random variables and $R \sim F$. Conversely, if (20),(21) holds with $R, \Theta_i, 1 \le i \le k$ independent random variables where $R \sim F$ is a positive random radius and the random angle Θ_i has the density function q_i defined above, then $\boldsymbol{X} \sim \mathcal{GSD}(k, p, \boldsymbol{\alpha}, F)$.

Remark 1. a) In view of Corollary 3, any subvector $\boldsymbol{X}_I, I \subset \{1, \ldots, k\}$ of a k-dimensional LpGSD random vector $\boldsymbol{X}$ has a LpGSD distribution function. Moreover $\boldsymbol{X}_I$ possesses a density function. This property was derived for elliptical distributions in Cambanis et al. (1981). Explicitly, let $\boldsymbol{X}_I$ be as defined in Corollary 3, then if follows from (19) that the associated random radius $R_{I,p}$ possesses the density function f given for any $u \in (0, \omega)$ by

$$f(u) = pu^{p\overline{\alpha}_I - 1} \frac{\Gamma(\overline{\alpha})}{\Gamma(\overline{\alpha}_I)\Gamma(\overline{\alpha} - \overline{\alpha}_I)} \int_u^\omega (r^p - u^p)^{\overline{\alpha} - \overline{\alpha}_I - 1} r^{-p(\overline{\alpha}-1)}\, dF(r). \tag{22}$$

Here ω is the upper endpoint of the distribution function F of R.

In the special case $\alpha_i = 1/p, i = 1, \ldots, k$, the expression for the p.d.f simplifies to (see e.g. Gupta and Song (1997))

$$f(u) = pu^{m-1} \frac{\Gamma(k/p)}{\Gamma(m/p)\Gamma((k-m)/p)} \int_u^\infty (r^p - u^p)^{(k-m)/p - 1} r^{-k+p}\, dF(r). \tag{23}$$

b) If $\alpha_i = i/p, 1 \le i \le k$, then Corollary 4 reduces to Theorem 2 in Szabłowski (1998). The case of L_2-norm spherical random vectors is presented in Theorem 2.11 of Fang et al. (1990).

Next we derive the conditional distribution $\boldsymbol{X}_I | \boldsymbol{X}_J = \boldsymbol{x}_J, \boldsymbol{x} \in \mathbb{R}^k$ where I, J are two non-empty disjoint index sets of $\{1, \ldots, k\}$. It follows that the conditional distribution is determined in terms of the norm $\|\boldsymbol{x}_J\|_p := (\sum_{j \in J} |x_j|^p)^{1/p}$.

We again emphasise that the results obtained in this section are similar to those for a much narrower classes of spherical and elliptical random vectors. Namely, the asymptotic results remain valid when in the basic stochastic representation (1), the L_2-norm uniformly distributed random vector $\boldsymbol{\mathcal{U}}$ is replaced by a L_p-norm generalised symmetrised Dirichlet random vector.

Theorem 5. *Let $\boldsymbol{X} \sim \mathcal{GSD}(k, p, \boldsymbol{\alpha}, F)$, with $p > 0, \boldsymbol{\alpha} \in (0, \infty)^k, k \ge 1$, and let I, J be partitions of $\{1, \ldots, k\}$. Then for any $\boldsymbol{x} \in \mathbb{R}^k$ with $F(\|\boldsymbol{x}_J\|_p) \in (0, 1)$ we have*

$$\boldsymbol{X}_I | \boldsymbol{X}_J = \boldsymbol{x}_J \stackrel{d}{=} R_{\|\boldsymbol{x}_J\|_p} \boldsymbol{\mathcal{V}}_I, \quad \boldsymbol{\mathcal{V}}_I \sim \mathcal{SD}(m, p, \boldsymbol{\alpha}_I). \tag{24}$$

Moreover $\boldsymbol{\mathcal{V}}_I$ is independent of $R_{\|\boldsymbol{x}_J\|_p} > 0$ with the distribution function G given by

$$G(x) := 1 - \frac{\int_{(\|\boldsymbol{x}_J\|_p^p + x^p)^{1/p}}^\omega (r^p - \|\boldsymbol{x}_J\|_p^p)^{\overline{\alpha}_I - 1} r^{-p\overline{\alpha}+p}\, dF(r)}{\int_{\|\boldsymbol{x}_J\|_p}^\omega (r^p - \|\boldsymbol{x}_J\|_p^p)^{\overline{\alpha}_I - 1} r^{-p\overline{\alpha}+p}\, dF(r)}, \quad \forall x > 0, \tag{25}$$

where $\omega \in (0, \infty]$ is the upper endpoint of the distribution function F.

4. Dependence and Asymptotic Dependence

A simple example of elliptical random vectors is $\boldsymbol{X} \sim N(\boldsymbol{\mu}, \Sigma)$ a Gaussian random vector in $\mathbb{R}^k, k \ge 2$, with the covariance matrix Σ and mean vector $\boldsymbol{\mu}$. It is well-known (see e.g. Fang et al. (1990)) that the independence of the components of $\boldsymbol{X}$ is equivalent to the assumption that Σ is the identity matrix. It is also

well-known (see e.g. Cambanis et al. (1981), Fang et al. (1990)) that a spherical random vector has independent components iff its components are Gaussian. Fang and Fang (1990) provide several conditions which imply the independence of the components of L_2-norm generalised symmetrised Dirichlet random vectors. In the next theorem we shall show that similar conditions are valid for a more general case of LpGSD random vectors.

Also it follows from the Theorem 6 that independence of components holds only in the case of Kotz Type I LpGSD distribution with parameters $N+1=s=1$ and $r>0$.

Theorem 6. *Let $\boldsymbol{X}$ be a L_p-norm generalised symmetrised Dirichlet random vector in $\mathbb{R}^k, k\geq 2$, with the density generator $g\sim\mathcal{G}(\boldsymbol{\alpha},\omega)$. The following statements are equivalent:*

(1) $\boldsymbol{X}$ possesses independent components.

(2) For any $I\subset\{1,\ldots,k\}$ the random vector $\boldsymbol{X}_I$ has Kotz Type I LpGSD distribution with parameters $N=0, s=1$ and $r>0$.

(3) There exist I,J disjoint index sets such that $\boldsymbol{X}_I$ is independent of $\boldsymbol{X}_J$.

(4) There exist I,J disjoint index sets with $I\cup J\subset\{1,\ldots,k\}$ such that $\boldsymbol{X}_I|\boldsymbol{X}_J$ is independent of $\boldsymbol{X}_J$.

(5) For any $I\subset\{1,\ldots,k\}$ we have $R^p_{I,p}\sim\Gamma(\overline{\alpha}_I,r)$ with $\overline{\alpha}_I=\sum_{i\in I}\alpha_i$ and r is a positive constant.

(6) There exist I,J disjoint index sets with $I\cup J\subset\{1,\ldots,k\}$ (provided $\overline{\alpha}_I\neq\overline{\alpha}_J$) such that the density generators of $\boldsymbol{X}_I$ and $\boldsymbol{X}_J$ differ only up to a positive constant.

The assumption of the above theorem that $\boldsymbol{X}$ possesses a density function is somewhat restrictive. In view of Theorem 2 any subvector $\boldsymbol{X}_I$, with $1\leq|I|<k$ possesses a density function even when $\boldsymbol{X}$ does not posses one. Thus if $k\geq 2$, the assumption that $\boldsymbol{X}$ has a density generator is not needed. Several statements given above could then be easily reformulated.

Next, we shall discuss the asymptotic dependence of LpGSD random vectors. Let $\boldsymbol{X}$ be as in Theorem 6 with the associated random radius $R\sim F$. A meaningful parameter for the asymptotic dependence between the components $X_i,X_j, 1\leq i<j<k$, is the limit (provided it exists)

$$\tau(X_i,X_j):=\lim_{t\uparrow\omega}\frac{\boldsymbol{P}\{X_i>t,X_j>t\}}{\boldsymbol{P}\{X_i>t\}+\boldsymbol{P}\{X_j>t\}},$$

where $\omega:=\sup\{x:F(x)<1\}$ is the upper endpoint of F.

If ω is finite then Theorem 2 implies that

$$\tau(X_i,X_j) \quad = \quad 0, \quad 1\leq i<j\leq k \tag{26}$$

since both X_i,X_j and $R_{I,p}, I=\{i,j\}$, have the same upper endpoint ω. Hence the joint tail probabilities diminish faster than each of the marginal tail probability.

The next result shows that (26) holds even if $\omega=\infty$, provided that the associated random radius R has a rapidly varying survival function $1-F$, i.e.

$$\lim_{t\to\infty}\frac{1-F(ct)}{1-F(t)} \quad = \quad 0 \tag{27}$$

for any $c>1$.

Theorem 7. *Let $\boldsymbol{X}$ be a LpGSD random vector in $\mathbb{R}^k, k\geq 2$, with the associated random radius R which is almost surely positive. If the distribution function F of*

R satisfies (27)*, then*

$$\tau(X_i, X_j/z) \quad = \quad 0, \quad 1 \le i < j \le k \tag{28}$$

is valid for any $z \in (0, \infty)$.

Example 8. [**Continue Example 4**]. Let $\boldsymbol{X} = R\boldsymbol{\mathcal{U}}$ be a k-dimensional Kotz Type III random vector. In view of the property (15) we have for any $c > 1$

$$\frac{\boldsymbol{P}\{R > cu\}}{\boldsymbol{P}\{R > u\}} \quad = \quad (1 + o(1))c^N \exp\big(-r[c^\delta - 1]u^\delta\big) \to 0, \quad u \to \infty.$$

This implies that the survival function $1 - F$ satisfies (27). Consequently (28) holds if $\boldsymbol{X}$ is a LpGSD random vector.

5. Conditional Limiting Theorems

Let $\boldsymbol{X}$ be as in Theorem 5 with $R \sim F$ such that F has the upper endpoint $\omega \in (0, \infty]$. Given I, J two subsets of $\{1, \ldots, k\}$ we shall derive in this section an asymptotic approximation for the distribution function of the conditional random vector $\boldsymbol{X}_I | \boldsymbol{X}_J = \boldsymbol{u}_J, \boldsymbol{u} \in \mathbb{R}^k$, letting $\boldsymbol{u}_J$ tend to some boundary point. Similar results for spherical and elliptical random vectors are derived in Hashorva (2006b,c,2007). In fact, the motivation for the aforementioned results comes from those previously reported in Berman (1992) where elliptical random vectors are discussed. As in Berman (1992) we assume a certain asymptotic tail behaviour of the distribution function F related to extreme value theory. Explicitly, we shall suppose that F is in the max-domain of attraction of an univariate extreme value distribution function H, i.e.

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} \Big| F^n(r(n)x + q(n)) - H(x) \Big| \quad = \quad 0, \tag{29}$$

where $r(n) > 0, q(n), n \ge 1$ are given constants.
We shall denote the above asymptotic relation by $F \in MDA(H)$, and refer the reader for a further insight in the extreme value theory to the following standard monographs: de Haan (1970), Leadbetter et al. (1983), Resnick (1987), Reiss (1989), Falk et al. (2004), Kotz and Nadarajah (2005).

We note in passing that H is either a) the unit Gumbel distribution $\Lambda(x) = \exp(-\exp(-x))$, or b) the unit Weibull distribution $\Psi_\gamma(x) = \exp(-|x|^\gamma), x < 0, \gamma > 0$, or c) the unit Fréchet distribution $\Phi_\gamma(x) = \exp(-x^{-\gamma}), x > 0, \gamma > 0$. The symbol ω denotes again the upper endpoint of the distribution function F.
We consider each case separately.

The Gumbel Case $F \in MDA(\Lambda)$:
If $H = \Lambda$ then (29) is equivalent to the fact that there exists a positive measurable function w such that

$$\lim_{u\uparrow\omega} \frac{1 - F(u + x/w(u))}{1 - F(u)} \quad = \quad \exp(-x), \quad \forall x \in \mathbb{R} \tag{30}$$

is valid. The positive scaling function w has the following asymptotic properties (see e.g. Resnick (1987) or Kotz and Nadarajah (2005))

$$\lim_{u\uparrow\omega} \frac{w(u + x/w(u))}{w(u)} \quad = \quad 1 \tag{31}$$

hold uniformly for x in compact sets of $\mathbb{R}$. Furthermore

$$\lim_{u\uparrow\omega} k(u)w(u) = \infty, \tag{32}$$

with $k(u) := u$ if $\omega = \infty$ and $k(u) := \omega - u$ otherwise.
It will be shown in the next theorem that the conditional distribution of LpGSD random vectors is approximated by a Kotz Type I LpGSD random vector, provided F satisfies the limiting condition (30). Evidently, the Kotz Type I LpGSD class of distributions includes the Gaussian distributions with correlation matrix equal the identity matrix. For L_2-norm spherical random vectors the limiting distribution is Gaussian (see Hashorva (2006b)). It is rather surprising that a large class of LpGDS distributions can be approximated by a distribution function (Kotz Type I), which is completely known, provided the associated random radius is in the Gumbel max-domain of attraction. Moreover, the limiting distribution has independent components!

Theorem 8. *Let $F, \boldsymbol{X}$ be as in Theorem 5 with $\omega \in (0,\infty]$ the upper endpoint of F, and I, J be two non-empty disjoint sets of $\{1,\ldots,k\}$. Assume that distribution function F is in the Gumbel max-domain of attraction with positive scaling function w. If $\boldsymbol{u}_n \in \mathbb{R}^k, n \geq 1$, is such that $\|\boldsymbol{u}_{n,J}\|_p < \omega, n \geq 1$, and furthermore*

$$\lim_{n\to\infty} \|\boldsymbol{u}_{n,J}\|_p \quad = \quad \omega, \tag{33}$$

we then have the convergence in the distribution

$$\left(\frac{w(\|\boldsymbol{u}_{n,J}\|_p)}{\|\boldsymbol{u}_{n,J}\|_p^{p-1}}\right)^{1/p} \Big(\boldsymbol{X}_I | \boldsymbol{X}_J = \boldsymbol{u}_{n,J}\Big) \quad \stackrel{d}{\to} \quad \boldsymbol{Z}, \quad n\to\infty, \tag{34}$$

where $\boldsymbol{Z} \sim \boldsymbol{\mathcal{K}}_{\boldsymbol{\alpha}_I,p}$ is a Kotz Type I LpGSD random vector in $\mathbb{R}^{|I|}$ with parameters $\boldsymbol{\alpha}_I, N+1 = r = s = 1$.

For $\boldsymbol{Z} \sim \boldsymbol{\mathcal{K}}_{\boldsymbol{\alpha}_I,p}$ we have the stochastic representation

$$\boldsymbol{Z} \stackrel{d}{=} \mathcal{R}_I \boldsymbol{\mathcal{V}}_I,$$

with $\mathcal{R}_I^p > 0$ independent of $\boldsymbol{\mathcal{V}}_I$ and moreover

$$\mathcal{R}_I^p \sim Gamma(\overline{\alpha}_I, 1/p), \quad \boldsymbol{\mathcal{V}}_I \sim \mathcal{SD}(|I|, p, \boldsymbol{\alpha}_I).$$

Consequently if $p = 2$ and $\boldsymbol{\alpha} = \mathbf{1}/2 \in \mathbb{R}^k$, then $\boldsymbol{Z}$ is a standard Gaussian random vector in $\mathbb{R}^{|I|}$ with independent components. The above theorem asserts that for a spherically distributed $\boldsymbol{X}$ the conditional limiting distribution is Gaussian with the identity correlation matrix. This is shown in Corollary 3.1 of Hashorva (2006b) which is motivated by Theorem 4.1 of Berman (1983) (see also Theorem 12.4.1 in Berman (1992) and Lemma 8.2 in Berman (1982)).
It is interesting to note that the Gaussian approximation of Type I spherical random vectors (L_2-norm) is a special case of the Kotz approximation of LpGSD Type I random vectors. We present next an example.

Example 9. [**Regularly varying scaling function**] Let $\boldsymbol{X}$ be a k-dimensional LpGSD random vector with associated random radius R which has distribution function F in the Gumbel max-domain of attraction with the scaling function

$$w(u) = (1+o(1))u^{\delta}L(u), \quad \delta > 0, \quad u\to\infty,$$

where L is a positive function such that $\lim_{u\to\infty} L(Ku)/L(u) = 1, \forall K > 1$.
Consider now positive constants $u_n, n \geq 1$ such that $\lim_{n\to\infty} u_n = \infty$ and let I, J be two non-empty disjoint subsets of $\{1, \ldots, k\}$. For a given vector $\boldsymbol{a} \in \mathbb{R}^k$ such that $\boldsymbol{a}_J \neq \boldsymbol{0}_J$ and any integer $n \geq 1$ define

$$\boldsymbol{u}_n := u_n \boldsymbol{a}, \quad h_n := \left(\frac{w(u_n \|\boldsymbol{a}_J\|_p)}{(u_n \|\boldsymbol{a}_J\|_p)^{p-1}} \right)^{1/p} = (L(u_n \|\boldsymbol{a}_J\|_p))^{1/p} (u_n \|\boldsymbol{a}_J\|_p)^{(\delta+1)/p-1}.$$

Clearly, $\boldsymbol{u}_n, n \geq 1$, satisfies (33), consequently Theorem 8 implies

$$h_n \boldsymbol{X}_I | \boldsymbol{X}_J = u_n \boldsymbol{a}_J \quad \stackrel{d}{\to} \quad \boldsymbol{Z} \sim \boldsymbol{\mathcal{K}}_{\boldsymbol{\alpha}_I, p}, \quad n \to \infty, \tag{35}$$

where $\boldsymbol{\mathcal{K}}_{\boldsymbol{\alpha}_I, p}$ denotes a standard Kotz Type I LpGSD random vector.

If $\boldsymbol{X}$ is a Kotz Type III as in Example 2, then the associated random radius has the distribution function in the Gumbel max-domain of attraction with the scaling function w given by

$$w(u) = (1 + o(1)) r\delta u^{\delta-1}, \quad u \to \infty.$$

This follows easily by observing that

$$\begin{aligned} \frac{\boldsymbol{P}\{R > u + x/w(u)\}}{\boldsymbol{P}\{R > u\}} &= (1 + o(1)) \Big(1 + \frac{x}{r\delta u^\delta}\Big)^N \exp\Big(-ru^\delta \Big[\Big(1 + \frac{x}{r\delta u^\delta}\Big)^\delta - 1\Big]\Big) \\ &\to \exp(-x), \quad u \to \infty. \end{aligned}$$

Hence for this case (35) holds with $h_n := (r\delta)^{1/p} (u_n \|\boldsymbol{a}_J\|_p)^{\delta/p-1}, n \geq 1$.

The Weibull Case $F \in MDA(\Psi_\gamma)$):
The distribution function F has necessarily a finite upper endpoint ω.
Without loss of generality we assume in the following theorem that $\omega = 1$. The conditional distribution of $\boldsymbol{X}_I | \boldsymbol{X}_J = \boldsymbol{u}_{n,J}$ can be approximated ($n \to \infty$) by another LpGSD random vector as shown in the next theorem.

Theorem 9. *Let $F, I, J, \boldsymbol{X}, \boldsymbol{u}_n, n \geq 1$ be as in Theorem 8 and $c_n, n \geq 1$, be a sequence of positive constants converging to 0. Assume that the upper endpoint of F is $\omega = 1$, and furthermore*

$$\lim_{n\to\infty} \frac{1 - \|\boldsymbol{u}_{n,J}\|_p}{c_n} = 1 \tag{36}$$

holds. If $F \in MDA(\Psi_\gamma), \gamma > 0$, we then have

$$\left(\frac{1}{pc_n}\right)^{1/p} \Big(\boldsymbol{X}_I | \boldsymbol{X}_J = \boldsymbol{u}_{n,J}\Big) \quad \stackrel{d}{\to} \quad \mathcal{R}_I \boldsymbol{\mathcal{V}}_I, \quad n \to \infty, \tag{37}$$

where $\mathcal{R}_I^p \sim Beta(\overline{\alpha}_I, \gamma + \overline{\alpha} - \overline{\alpha}_I - \overline{\alpha}_J), \mathcal{R}_I > 0$ and $\mathcal{R}_I$ is independent of $\boldsymbol{\mathcal{V}}_I \sim \mathcal{SD}(|I|, p, \boldsymbol{\alpha}_I)$.

We note in passing that Theorem 12.7.1. in Berman (1992) a related result to (37) is shown for a bivariate elliptical random vector. The multivariate extension of Berman's theorem is presented in Theorem 3.2 of Hashorva (2007).

Example 10. [**Kummer-Beta**] Let $\boldsymbol{X}$ be as in Example 6. It follows that the random radius R associated with $\boldsymbol{X}$ has the distribution function in the Weibull max-domain of attraction of Ψ_γ. Consider the sequence of vectors $\boldsymbol{u}_n = (1 -$

$1/n, 0, \ldots, 0), n \geq 1$ in $\mathbb{R}^k$. If $I = \{1, \ldots, r\}, 1 \leq r < k$, then (36) holds with $c_n = 1/n, n \geq 1$. Consequently Theorem 9 implies the convergence

$$\left(\frac{n}{p}\right)^{1/p} \Big(\boldsymbol{X}_I | \boldsymbol{X}_J = \boldsymbol{u}_{n,J}\Big) \quad \stackrel{d}{\to} \quad \mathcal{R}_I \boldsymbol{\mathcal{V}}_I, \quad n \to \infty,$$

where $J = \{r+1, \ldots, m\}, r < m \leq k, \mathcal{R}_I \sim Beta(\sum_{i=1}^r \alpha_i, \gamma + k - m)$ and $\boldsymbol{\mathcal{V}}_I \sim \mathcal{SD}(r, p, \boldsymbol{\alpha}_I)$.

The Fréchet Case $F \in MDA(\Phi_\gamma)$:
In this case F has an infinite upper endpoint. Similarly to the two other cases of max-domain of attraction it is possible also to approximate here the conditional distribution of LpGSD random vectors. We have:

Theorem 10. *Let $F, I, J, \boldsymbol{X}, c_n, \boldsymbol{u}_n, n \geq 1$ be as in Theorem 9 such that*

$$\lim_{n\to\infty} c_n \boldsymbol{u}_{n,J} = \boldsymbol{u}_J \neq \boldsymbol{0}_J \tag{38}$$

is valid. If $F \in MDA(\Phi_\gamma), \gamma > 0$, we then have the convergence in the distribution

$$c_n \Big(\boldsymbol{X}_I | \boldsymbol{X}_J = \boldsymbol{u}_{n,J}\Big) \stackrel{d}{\to} \boldsymbol{Y}_I | \boldsymbol{Y}_J = \boldsymbol{u}_J, \quad n \to \infty, \tag{39}$$

where $\boldsymbol{Y} \sim \mathcal{GSD}(k, p, \boldsymbol{\alpha}, F_\gamma)$ with the distribution function F_γ defined by

$$F_\gamma(r) := 1 - (r/\|\boldsymbol{u}_J\|_p)^{-\gamma - p(\overline{\alpha} - \overline{\alpha}_I - \overline{\alpha}_J)}, \quad \forall r \geq \|\boldsymbol{u}_J\|_p.$$

A natural choice for the constants $c_n, n \geq 1$ in the above theorem is $c_n := 1/\|\boldsymbol{u}_{n,J}\|_p, n \geq 1$, provided that $\lim_{n\to\infty} \|\boldsymbol{u}_{n,J}\|_p = \infty$. The latter is actually a necessary condition for (38) to hold.

Example 11. [**Kotz Type III**] Let $\boldsymbol{X}$ be as in Example 4 with $N < 0, \delta \leq 0, p > 0$. Then the associated random radius R of $\boldsymbol{X}$ has the distribution function in the max-domain of attraction of the Fréchet distribution Φ_{-N}. Consequently, if $\boldsymbol{X}$ is also a LpGSD random vector, Theorem 10 implies for I, J disjoint index sets and $u_n > 0, n \geq 1$ such that $\lim_{n\to\infty} u_n = \infty$

$$\frac{1}{u_n} \Big(\boldsymbol{X}_I | \boldsymbol{X}_J = u_n \boldsymbol{u}_J\Big) \stackrel{d}{\to} \boldsymbol{Y}_I | \boldsymbol{Y}_J = \boldsymbol{u}_J, \quad n \to \infty,$$

where $\|\boldsymbol{u}_J\|_p > 0$ and $\boldsymbol{Y} \sim \mathcal{GSD}(k, p, \boldsymbol{\alpha}, F_{-N})$, with the distribution function F_{-N} given by

$$F_{-N}(r) = 1 - (r/\|\boldsymbol{u}_J\|_p)^{N - p(k-m)}, \quad \forall r \geq \|\boldsymbol{u}_J\|_p.$$

6. Tail Asymptotics

Let $\boldsymbol{X}$ be a k-dimensional LpGDS random vector with the associated random radius R. The distributional properties of $\boldsymbol{X}$ are determined by those of R. Similarly we expect that in an asymptotic context the asymptotic behaviour of $\boldsymbol{P}\{\boldsymbol{X}/n \in B\}, n \to \infty$, with B being a Borel set, is defined by the tail asymptotics of R. In the special cases when R has distribution function in the max-domain of attraction of an univariate extreme value distribution H, then the tail asymptotic behaviour of X_1 can be determined by applying Lemma 16 in the Appendix.
The case where R is regularly varying with index $\gamma \geq 0$, i.e.

$$\lim_{t\to\infty} \frac{\boldsymbol{P}\{R > tx\}}{\boldsymbol{P}\{R > t\}} \quad = \quad x^{-\gamma}, \quad \forall x > 0,$$

is quite tractable as shown in Hashorva (2006a). The above asymptotic relation defines the tail asymptotic of $\boldsymbol{X}$ and in particular of its components, and moreover the converse is true.

Indeed we have the following result:

Theorem 11. *Let $\boldsymbol{X} \sim \mathcal{GSD}(k, p, \boldsymbol{\alpha}, F)$ be a k-dimensional random vector with the associated random radius R, and $A \in \mathbb{R}^{k\times k}$ be a non-singular matrix. The statements below are then equivalent:*

i) $|X_1|$ is regularly varying with a positive index γ.

ii) For any non-empty $I \subset \{1, \ldots, k\}$ the random radius $\|\boldsymbol{X}_I\|_p$ is regularly varying with index $\gamma > 0$, and furthermore if $|I| < k$ then

$$\boldsymbol{P}\{\|\boldsymbol{X}_I\|_p > u\} = (1+o(1))\frac{\Gamma(\overline{\alpha})\Gamma(\overline{\alpha}_I+\gamma/p)}{\Gamma(\overline{\alpha}_I)\Gamma(\overline{\alpha}+\gamma/p)}\boldsymbol{P}\{\|\boldsymbol{X}\|_p > u\}, \quad u\to\infty. \tag{40}$$

iii) For any non-empty $I \subset \{1, \ldots, k\}$ with m elements and any Borel set $B \subset \mathbb{R}^m$ not containing the origin $\mathbf{0} \in \mathbb{R}^m$

$$\begin{aligned}&\lim_{u\to\infty}\frac{\boldsymbol{P}\{(A_{II}\boldsymbol{X}_I+\boldsymbol{\mu}_I)/u \in B\}}{\boldsymbol{P}\{X_i > u\}}\\ &= \frac{2\gamma\Gamma(\alpha_i)\Gamma(\overline{\alpha}+\gamma/p)}{\Gamma(\overline{\alpha})\Gamma(\alpha_i+\gamma/p)}\int_0^\infty \boldsymbol{P}\{rA_{II}\boldsymbol{\mathcal{V}}_I \in B\}r^{-\gamma-1}\,dr\end{aligned} \tag{41}$$

holds with $i \le k, \boldsymbol{\mu} \in \mathbb{R}^k$ and $\boldsymbol{\mathcal{V}}_I \sim \mathcal{SD}(m, p, \boldsymbol{\alpha}_I)$.

Corollary 12. *Let $A, R, \boldsymbol{X}$ be as in Theorem 11. Assume that the associated random radius R or the first components X_1 of $\boldsymbol{X}$ is regularly varying with positive index γ. Let $X_i \sim G_i, i \le k, a_i(n) := G_i^{-1}(1-1/n), n > 1$ with G_i^{-1} the generalised inverse of G_i and set $\boldsymbol{Y} := A\boldsymbol{X}$. Then we have for any $\boldsymbol{y} = (y_1, \ldots, y_k)^\top > \mathbf{0}$*

$$\begin{aligned}&\lim_{n\to\infty}\boldsymbol{P}\{Y_1/a_1(n) \le y_1, \ldots, Y_k/a_k(n) \le y_k\}^n\\ &= \exp\Big(-\gamma\int_0^\infty \boldsymbol{P}\{r\boldsymbol{c}A\boldsymbol{\mathcal{U}} \not\le \boldsymbol{y}\}r^{-\gamma-1}\,dr\Big),\end{aligned} \tag{42}$$

where the vector $\boldsymbol{c} = (c_1, \ldots, c_k)^\top$ has components given by

$$c_i := \left(\frac{\Gamma(\overline{\alpha})\Gamma(\alpha_i+\gamma/p}{2\Gamma(\alpha_i)\Gamma(\overline{\alpha}+\gamma/p)}\right)^{-1/\gamma}, \quad i \le k.$$

Remark 2. i) Statement *iii*) in Theorem 11 implies that $\boldsymbol{X}$ is a multivariate regularly varying random vector in $\mathbb{R}^k$ with index $\gamma > 0$, and in particular $|X_i|, i \le k$ is regularly varying with index γ. See Basrak at al. (2002) for more details on regular variation of random vectors.

ii) If A is the identity matrix then the right-hand side of (42) is a distribution function with the Fréchet marginal distributions $\Phi_\gamma(x) = \exp(-x^{-\gamma}), x > 0$.

iii) Using (41) we obtain for any $\delta, \lambda, p, \gamma$ positive with $\lambda - \delta > 0$

$$\frac{\gamma\Gamma(\delta)\Gamma(\lambda+\gamma/p}{\Gamma(\lambda)\Gamma(\delta+\gamma/p)}\int_1^\infty \boldsymbol{P}\{Z > r^{-p}\}r^{-\gamma-1}\,dr = 1,$$

with $Z \sim Beta(\delta, \lambda-\delta)$. Consequently we have

$$h_{\delta,\lambda,\gamma}(x) := \frac{\gamma\Gamma(\delta)\Gamma(\lambda+\gamma/p)}{\Gamma(\lambda)\Gamma(\delta+\gamma/p)}\int_x^\infty \boldsymbol{P}\{Z > r^{-p}\}r^{-\gamma-1}\,dr, \quad x \ge 1$$

is a survival function of a positive random variable in $[1, \infty)$.

Example 12. [**Kotz Type II**] Let as in Example 3 $\boldsymbol{X}$ be a Kotz Type II LpGSD random vector with parameters $\boldsymbol{\alpha}, s < 0, N < -\overline{\alpha}$. It follows that R is regularly varying with the index $-p(N+\overline{\alpha})$. Hence, any component of $\boldsymbol{X}$ is regularly varying with the index $-p(N+\overline{\alpha})$.

7. Proofs

Proof of Theorem 1: We carry out the following transformations of variables

$$y_i = x_i r^{-p}, \quad i = 1, \ldots, k-1, \text{ and } r^p = \sum_{i=1}^{k} |x_i|^p.$$

Calculating the Jacobian of this transformation we arrive at the p.d.f of $\boldsymbol{X}$

$$\begin{aligned} h(r, z_1, \ldots, z_{k_1}) &= 2g(r^p) r^{p\sum_{i=1}^k \alpha_i - 1} \prod_{i=1}^{k-1} |z_i|^{p\alpha_i - 1} \Big(1 - \sum_{i=1}^{k-1} |z_i|^p\Big)^{\alpha_k - 1} \\ &= \frac{2^k \prod_{i=1}^k \Gamma(\alpha_i)}{p^{k-1}\Gamma(\overline{\alpha})} g(r^p) r^{p\sum_{i=1}^k \alpha_i - 1} c \prod_{i=1}^{k-1} |z_i|^{p\alpha_i - 1} \Big(1 - \sum_{i=1}^{k-1} |z_i|^p\Big)^{\alpha_k - 1}, \end{aligned}$$

with $c^{-1} := (2/p)^{k-1} \prod_{i=1}^k \Gamma(\alpha_i)/\Gamma(\overline{\alpha})$. Hence the result follows by recalling the form of density function in (8).
Now, if $\boldsymbol{X}$ has p.d.f h given by (6), then in view of (1) the p.d.f of $R\boldsymbol{\mathcal{U}}$ is given by

$$h(r, u_1, \ldots, u_{k-1}) = f(r) q(u_1, \ldots, u_{k-1}),$$

with q as in (2). Transforming the variables as above it follows that $\boldsymbol{X}$ has p.d.f h given by (6), hence the proof is complete.

□

Proof of Theorem 2: The proof is analogous (considering p instead of 2) to the proof of Theorem 4.1 of Fang and Fang (1990). For the sake of completeness we shall provide a sketch. First note that $\boldsymbol{X} \stackrel{d}{=} R\boldsymbol{\mathcal{U}}$ with R independent of $\boldsymbol{\mathcal{U}} \sim \mathcal{SD}(k, p, \boldsymbol{\alpha})$. The properties of $\boldsymbol{\mathcal{U}}$ and in particular can be derived considering $\boldsymbol{X} \sim \boldsymbol{\mathcal{K}}_{\boldsymbol{\alpha},p}$ as in Example 2 ($N + 1 = r = s = 1$). Since also $\boldsymbol{X}_{I_j}, 1 \le j \le m$ are LpGSD random vectors we have

$$\boldsymbol{X}_{I_j} \stackrel{d}{=} R_{I_j} \boldsymbol{\mathcal{V}}_{I_j}, \quad 1 \le j \le m,$$

with $R_{I_j} \stackrel{d}{=} \|\boldsymbol{X}_{I_j}\|_p$ independent of $\boldsymbol{\mathcal{V}}_{I_j} \sim \mathcal{SD}(|I_j|, p, \boldsymbol{\alpha}_{I_j})$.

By Lemma 13 and (13) we have

$$R^p \stackrel{d}{=} \|\boldsymbol{X}\|_p^p = \Gamma(\overline{\alpha}, 1/p), \quad R_{I_j}^p \stackrel{d}{=} \|\boldsymbol{X}_{I_j}\|_p^p = \Gamma(a_j, 1/p), \quad 1 \le j \le m,$$

with $a_j := \sum_{i \in I_j} \alpha_i$. Denote $\boldsymbol{a}_m := (a_1, \ldots, a_m)^\top \in (0, \infty)^m$ and

$$\boldsymbol{Z}_j := \frac{\boldsymbol{X}_{I_j}}{\|\boldsymbol{X}_{I_j}\|_p}, \quad \boldsymbol{W}_m := (W_1, \ldots, W_m)^\top, \text{ with } W_j := \frac{\|\boldsymbol{X}_{I_j}\|_p}{\|\boldsymbol{X}\|_p}, 1 \le j \le m.$$

Evidently, $\|\boldsymbol{W}\|_p = 1$ almost surely. The proof now follows easily since $\boldsymbol{W}_m \sim \mathcal{SD}(m, p, \boldsymbol{a}_m)$.

□

PROOF OF COROLLARY 3: Let $J = \{1, \ldots, k\} \setminus I$, and F denote the distribution function of RW where $W > 0$ is independent of R with $W \sim Beta(\overline{\alpha}_I, \overline{\alpha}_J)$. Applying Theorem 2 to partitions I, J we obtain

$$\boldsymbol{X}_I \stackrel{d}{=} R_{I,p}\boldsymbol{\mathcal{V}}_I, \text{ with } \boldsymbol{\mathcal{V}}_I \sim \mathcal{SD}(|I|, p, \boldsymbol{\alpha}_I).$$

Since $R_{I,p}$ independent of $\boldsymbol{\mathcal{U}}$, $\boldsymbol{X}_I$ is a LpGSD random vector in $\mathbb{R}^{|I|}$. Using Lemma 13 we have

$$R_{I,p} = \|\boldsymbol{X}_I\|_p \stackrel{d}{=} \|RW\boldsymbol{\mathcal{V}}_I\|_p = RW\|\boldsymbol{\mathcal{V}}_I\|_p = RW,$$

and the proof is completed.

□

PROOF OF COROLLARY 4: Let $\mathcal{I}_i, i \leq k$ be independent random variables taking values $-1, 1$ with probability $1/2$. For simplicity we show the proof when $k = 3$. The general case $k > 3$ follows utilising similar arguments. By the assumption $\boldsymbol{X} \stackrel{d}{=} R\boldsymbol{\mathcal{U}}$, with $\boldsymbol{\mathcal{U}} \sim \mathcal{SD}(3, p, (\alpha_1, \alpha_2, \alpha_3))$ independent of R. In view of Theorem 2, we have

$$\boldsymbol{X} \stackrel{d}{=} R\Big(\mathcal{I}_1(1 - V_1^p)^{1/p}, \mathcal{I}_2 V_1(1 - V_2^p)^{1/p}, \mathcal{I}_3 V_1 V_2\Big),$$

where $V_1^p \sim Beta(\alpha_2 + \alpha_3, \alpha_1), V_2^p \sim Beta(\alpha_3, \alpha_2)$. Define the random angles $\Theta_1 \in [0, \pi], \Theta_2 \in [0, 2\pi]$ such that

$$\sin(\Theta_1) := V_1^{p/2}, \quad |\cos(\Theta_1)|^{2/p} := (1 - V_1^p)^{1/p},$$
$$|\cos(\Theta_2)|^{2/p} := (1 - V_2^p)^{1/p}, \quad |\sin(\Theta_2)| := V_2^{p/2},$$

and $\text{sign}(\cos(\Theta_2))$ independent of $\text{sign}(\sin(\Theta_2))$ two symmetric random variables. It follows that the density function of Θ_1 is given by

$$q_1(\theta) := \frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)}{\Gamma(\alpha_2 + \alpha_3)\Gamma(\alpha_1)} |\sin(\theta)|^{2(\alpha_2+\alpha_3)-1} |\cos(\theta)|^{2\alpha_1 - 1}, \quad \theta \in [0, \pi],$$

and Θ_2 has density function

$$q_2(\theta) := \frac{1}{2} \frac{\Gamma(\alpha_2 + \alpha_3)}{\Gamma(\alpha_3)\Gamma(\alpha_2)} |\sin(\theta)|^{2\alpha_3 - 1} |\cos(\theta)|^{2\alpha_2 - 1}, \quad \theta \in [0, 2\pi].$$

Hence we have the stochastic representation

$$X_1 \stackrel{d}{=} R|\cos(\Theta_1)|^{2/p} \text{sign}(\cos(\Theta_1)), \; X_2 \stackrel{d}{=} R[\sin(\Theta_1)|\cos(\Theta_2)|]^{2/p} \text{sign}(\cos(\Theta_2)),$$
$$X_3 \stackrel{d}{=} R[\sin(\Theta_1)|\sin(\Theta_2)|]^{2/p} \text{sign}(\sin(\Theta_2)).$$

The converse follows easily by reversing the argument.

□

PROOF OF THEOREM 5: The proof is based on the stochastic representation (1) and the amalgamation property with respect to the partitions I, J. It follows along the lines of the proof of Theorem 5 in Cambanis et al. (1981).

□

PROOF OF THEOREM 6: The amalgamation property of LpGSD random vectors shows that the conditional and marginal distributions of L_p-norm generalised symmetrised Dirichlet random vectors are of the same form as those for the case of L_2-norm. Thus the proof of the general case $p > 0$ follows by utilising the same arguments as in the proof of Theorem 4.3 in Fang and Fang (1990).

□

PROOF OF THEOREM 7: Let $i, j, i \neq j$ be given and $z > 0, c_0 \in (0,1)$ be constants. Set $k_z^p := \inf\{|x_1|^p + |x_2|^p : x_1 \geq 1, x_2 \geq z\} \geq 1$ which does exist. In view of Corollary 3, we obtain for any $t \in (0, \omega)$ (write $I := \{i, j\}$)

$$\begin{aligned}\frac{\boldsymbol{P}\{X_i > t, X_j > tz\}}{\boldsymbol{P}\{X_i > t\} + \boldsymbol{P}\{X_j > tz\}} &\leq \frac{\boldsymbol{P}\{|X_i|^p + |X_j|^p \geq k_z t^p\}}{\boldsymbol{P}\{X_i > t\}} \\ &= \frac{2\boldsymbol{P}\{R_{I,p} \geq k_z t\}}{\boldsymbol{P}\{|X_i| > t\}} \\ &\leq \frac{2\boldsymbol{P}\{R \geq k_z t\}}{\boldsymbol{P}\{RW_i > t, W_1 > c_0\}} \\ &\leq \frac{2\boldsymbol{P}\{R \geq k_z t\}}{\boldsymbol{P}\{R > t/c_0\}\boldsymbol{P}\{W_i > c_0\}},\end{aligned}$$

with $W_i > 0$ almost surely such that $W_i^p \sim Beta(\alpha_i, \overline{\alpha} - \alpha_i)$. Since the survival function of the random radius R is rapidly varying and $k_z \geq 1$, the claim follows by choosing $c_0 \in (1/k_z, 1)$ and then letting $t \to \infty$.

□

PROOF OF THEOREM 8: Let $\omega \in (0, \infty]$ denote the upper endpoint of the distribution function F and set

$$a_n := \|\boldsymbol{u}_{n,J}\|_p, \quad \text{and } w_n := w(a_n), \qquad n \geq 1.$$

By the assumptions $\lim_{n\to\infty} a_n = \omega$ and $F \in MDA(\Lambda)$ with a positive scaling function w, we thus obtain that in view of (31) and (32)

$$\lim_{n\to\infty} a_n w_n = \infty, \quad \lim_{n\to\infty} w_n(\omega - a_n) = \infty. \tag{43}$$

Theorem 5 and Lemma 16 imply that $\boldsymbol{X}_{I\cup J}$ is a *LpGSD* random vector with the associated random radius R^* which has distribution function in the max-domain of attraction of Λ and the scaling function w. Therefore, we may assume for simplicity $I \cup J = \{1, \ldots, k\}$. In view of Theorem 5, we have for any large n

$$\Big(\boldsymbol{X}_I | \boldsymbol{X}_J = \boldsymbol{u}_{n,J}\Big) \stackrel{d}{=} R_{a_n} \boldsymbol{\mathcal{V}}_I, \quad \boldsymbol{\mathcal{V}}_I \sim \mathcal{SD}(|I|, p, \boldsymbol{\alpha}_I), \tag{44}$$

with $\boldsymbol{\mathcal{V}}_I$ independent of R_{a_n} such that

$$\boldsymbol{P}\{R_{a_n} > x\} = \frac{\int_{(a_n^p + x^p)^{1/p}}^{\omega} (r^p - a_n^p)^{\overline{\alpha}_I - 1} r^{-p\overline{\alpha}+p}\, dF(r)}{\int_{a_n}^{\omega} (r^p - a_n^p)^{\overline{\alpha}_I - 1} r^{-p\overline{\alpha}+p}\, dF(r)}$$

for all $x \in (0, (\omega^p - a_n^p)^{1/p})$. Furthermore, (30) implies that for any $s \in \mathbb{R}$

$$\lim_{n\to\infty} \frac{1 - F(a_n + s/w_n)}{1 - F(a_n)} = \exp(-s).$$

Hence the sequence of distribution functions

$$F_n(s) := \frac{F(a_n + s/w_n) - F(a_n)}{1 - F(a_n)}, \quad s \geq 0, n \geq 1$$

converges uniformly to the unit exponential distribution as $n \to \infty$. Moreover (43) implies for any $x \geq 0$

$$(x^p a_n^{p-1}/w_n + a_n^p)^{1/p} = a_n + (1 + o(1))p^{-1}x^p/w_n, \quad n \to \infty.$$

Transforming the variables we have for n large

$$\begin{aligned}
&\boldsymbol{P}\{R_{a_n} > x(a_n^{p-1}/w_n)^{1/p}\} \\
&= \frac{\int_{a_n+(1+o(1))p^{-1}x^p/w_n}^{\omega}(r^p - a_n^p)^{\overline{\alpha}_I-1}r^{-p\overline{\alpha}+p}\,dF(r)}{\int_{a_n}^{\omega}(r^p - a_n^p)^{\overline{\alpha}_I-1}r^{-p\overline{\alpha}+p}\,dF(r)} \\
&= \frac{\int_{x^p/p}^{w_n(\omega-a_n)}((a_n+s/w_n)^p - a_n^p)^{\overline{\alpha}_I-1}(a_n+s/w_n)^{-p\overline{\alpha}+p}\,dF_n(s)}{\int_0^{w_n(\omega-a_n)}((a_n+s/w_n)^p - a_n^p)^{\overline{\alpha}_I-1}(a_n+s/w_n)\,dF_n(s)} \\
&= \frac{\int_{x^p/p}^{w_n(\omega-a_n)} s^{\overline{\alpha}_I-1}(1+o(1))\,dF_n(s)}{\int_0^{w_n(\omega-a_n)} s^{\overline{\alpha}_I-1}(1+o(1))\,dF_n(s)}, \quad n\to\infty.
\end{aligned}$$

Lemma 4.4 of Hashorva (2006b) implies that

$$\lim_{n\to\infty}\int_{x^p/p}^{w_n(\omega-a_n)} s^{\overline{\alpha}_I-1}\,dF_n(s) = \int_{x^p/p}^{\infty} s^{\overline{\alpha}_I-1}\exp(-s)\,ds, \quad \forall x\geq 0.$$

Hence we obtain for any $x>0$

$$\lim_{n\to\infty}\boldsymbol{P}\{R_{a_n} > x(a_n^{p-1}/w_n)^{1/p}\} = \frac{1}{\Gamma(\overline{\alpha}_I)}\int_{x^p/p}^{\infty} s^{\overline{\alpha}_I-1}\exp(-s)\,ds =: \boldsymbol{P}\{\mathcal{R}_I > x\}.$$

Consequently we have the convergence in distribution

$$\left(\frac{w_n}{a_n^{p-1}}\right)^{1/p} R_{a_n} \quad \stackrel{d}{\to} \quad \mathcal{R}_I, \quad n\to\infty,$$

where $\mathcal{R}_I>0$ such that $\mathcal{R}_I^p \sim \Gamma(\overline{\alpha}_I, 1/p)$. Noting that $\mathcal{R}_I$ is independent of $\boldsymbol{\mathcal{U}}$ we arrive at the desired result.

□

Proof of Theorem 9: Let $a_n, R_{a_n}, n\geq 1$ be as in the proof of Theorem 8 and set $F_n(s) := F(1-c_n s), s\geq 0, n\geq 1$. For simplicity we assume that $a_n = 1-c_n, n\geq 1$ and denote $h_n := (1-F(a_n))^{-1}, n\geq 1$. If $I\cup J$ has less then k elements then Theorem 5 and Lemma 16 (Appendix) imply that the random vector $\boldsymbol{X}_{I\cup J}$ is a LpGSD random vector with associated random radius in the max-domain of attraction of Weibull distribution $\Psi_{\gamma^*}, \gamma^* := \gamma+\overline{\alpha}-\overline{\alpha}_I-\overline{\alpha}_J>0$. For simplicity, we consider therefore the case that $I\cup J = \{1,\ldots,k\}$ only. Since $F\in MDA(\Psi_\gamma)$ we have (see Kotz and Nadarajah (2005))

$$\lim_{n\to\infty} h_n[1-F_n(s)] \quad = \quad s^\gamma, \quad \forall s>0.$$

Furthermore $\lim_{n\to\infty} c_n = 0$ implies that

$$(pc_nx^p + (1-c_n)^p)^{1/p} \quad = \quad 1 - c_n(1-x^p)(1+o(1)), \quad n\to\infty.$$

Transforming the variables we obtain

$$\begin{aligned}
&\boldsymbol{P}\{R_{a_n} > x(pc_n)^{1/p}\} \\
&= \frac{\int_{(pc_nx^p+(1-c_n)^p)^{1/p}}^1 (r^p-(1-c_n)^p)^{\overline{\alpha}_I-1} r^{-p\overline{\alpha}+p}\, dF(r)}{\int_{1-c_n}^1 (r^p-(1-c_n)^p)^{\overline{\alpha}_I-1} r^{-p\overline{\alpha}+p}\, dF(r)} \\
&= \frac{\int_{1-c_n(1-x^p)(1+o(1))}^1 (r^p-(1-c_n)^p)^{\overline{\alpha}_I-1} r^{-p\overline{\alpha}+p}\, dF(r)}{\int_{1-c_n}^1 (r^p-(1-c_n)^p)^{\overline{\alpha}_I-1} r^{-p\overline{\alpha}+p}\, dF(r)} \\
&= \frac{\int_0^{(1-x^p)(1+o(1))} (1-s)^{\overline{\alpha}_I-1}(1+o(1))\, d(h_nF_n(s))}{\int_0^1 (1-s)^{\overline{\alpha}_I-1}(1+o(1))\, d(h_nF_n(s))}.
\end{aligned}$$

Utilising similar arguments as in Theorem 3.2 in Hashorva (2007) we have

$$\begin{aligned}
\lim_{n\to\infty} \boldsymbol{P}\{R_{a_n} > x(pc_n)^{1/p}\} &= \frac{\int_0^{1-x^p} (1-s)^{\overline{\alpha}_I-1} s^{\gamma-1}\, ds}{\int_0^1 (1-s)^{\overline{\alpha}_I-1} s^{\gamma-1}\, ds} \\
&= 1 - \frac{\int_0^{x^p} s^{\overline{\alpha}_I-1}(1-s)^{\gamma-1}\, ds}{\int_0^1 s^{\overline{\alpha}_I-1}(1-s)^{\gamma-1}\, ds}.
\end{aligned}$$

We thus have the convergence in the distribution

$$\left(\frac{1}{pc_n}\right)^{1/p} R_{a_n} \quad \stackrel{d}{\to} \quad \mathcal{R}_I, \quad n\to\infty,$$

where $\mathcal{R}_I$ satisfies $\mathcal{R}_I^p \sim Beta(\overline{\alpha}_I, \gamma)$ almost surely. Now, the proof follows using (44) and the fact that $\mathcal{R}_I$ is independent of $\boldsymbol{\mathcal{U}}$.

□

Proof of Theorem 10: Let $a_n, R_{a_n}, n \geq 1$, be as in the proof of Theorem 8 and set

$$F_\gamma(r) := 1 - \left(\frac{r}{\|\boldsymbol{u}_J\|_p}\right)^{-\gamma-p(\overline{\alpha}-\overline{\alpha}_I-\overline{\alpha}_J)}, \quad \forall r \geq \|\boldsymbol{u}_J\|_p.$$

Evidently, F_γ is a distribution function on $[\|\boldsymbol{u}_J\|_p, \infty)$. In view of Theorem 5 and Lemma 16 we need to show the claim only for the case $I \cup J = \{1,\ldots,k\}$. Assume now for simplicity that $\boldsymbol{u}_{n,J} = d_n\boldsymbol{u}_J, d_n := 1/c_n, n \geq 1$. The upper endpoint of the distribution function F is ∞ by the assumption. For any $x > 0$ and large n we have

$$\begin{aligned}
\boldsymbol{P}\{R_{a_n} > d_nx\} &= \frac{\int_{d_n(\|\boldsymbol{u}_J\|_p^p+x^p)^{1/p}}^\infty (r^p - d_n^p\|\boldsymbol{u}_J\|_p^p)^{\overline{\alpha}_I-1} r^{-p\overline{\alpha}+p}\, dF(r)}{\int_{d_n\|\boldsymbol{u}_J\|_p}^\infty (r^p - d_n^p\|\boldsymbol{u}_J\|_p^p)^{\overline{\alpha}_I-1} r^{-p\overline{\alpha}+p}\, dF(r)} \\
&= \frac{\int_{(\|\boldsymbol{u}_J\|_p^p+x^p)^{1/p}}^\infty (r^p - \|\boldsymbol{u}_J\|_p^p)^{\overline{\alpha}_I-1} r^{-p\overline{\alpha}+p}\, d(h_nF(d_nr))}{\int_{\|\boldsymbol{u}_J\|_p}^\infty (r^p - \|\boldsymbol{u}_J\|_p^p)^{\overline{\alpha}_I-1} r^{-p\overline{\alpha}+p}\, d(h_nF(d_nr))},
\end{aligned}$$

with $h_n := (1-F(d_n))^{-1}, n \geq 1$. Since $\lim_{n\to\infty} d_n = \infty$, the assumption on F implies

$$\lim_{n\to\infty} h_n[1-F(d_nx)] \quad = \quad x^{-\gamma}, \quad \forall x > 0,$$

hence we have by Fatou Lemma (see e.g. Kallenberg (1997))

$$\begin{aligned}\liminf_{n\to\infty}\int_{(\|\boldsymbol{u}_J\|_p^p+x^p)^{1/p}}^{\infty}&(r^p-\|\boldsymbol{u}_J\|_p^p)^{\overline{\alpha}_I-1}r^{-p\overline{\alpha}+p}\,d(h_nF(d_nr))\\ &\geq\int_{(\|\boldsymbol{u}_J\|_p^p+x^p)^{1/p}}^{\infty}(r^p-\|\boldsymbol{u}_J\|_p^p)^{\overline{\alpha}_I-1}r^{-p\overline{\alpha}+p}\,d(r^{-\gamma}).\end{aligned}$$

It follows by the Karamata Theorem (see e.g. Resnick (1987))

$$\begin{aligned}\limsup_{n\to\infty}\int_{(\|\boldsymbol{u}_J\|_p^p+x^p)^{1/p}}^{\infty}&(r^p-\|\boldsymbol{u}_J\|_p^p)^{\overline{\alpha}_I-1}r^{-p\overline{\alpha}+p}\,d(h_nF(d_nr))\\ &\leq\int_{(\|\boldsymbol{u}_J\|_p^p+x^p)^{1/p}}^{\infty}(r^p-\|\boldsymbol{u}_J\|_p^p)^{\overline{\alpha}_I-1}r^{-p\overline{\alpha}+p}\,d(r^{-\gamma}).\end{aligned}$$

Consequently

$$\begin{aligned}&\lim_{n\to\infty}\boldsymbol{P}\{R_{a_n}>d_nx\}\\ &=\frac{\int_{(\|\boldsymbol{u}_J\|_p^p+x^p)^{1/p}}^{\infty}(r^p-\|\boldsymbol{u}_J\|_p^p)^{\overline{\alpha}_I-1}r^{-p(\overline{\alpha}_I+\overline{\alpha}_J)+p}r^{-p(\overline{\alpha}-\overline{\alpha}_I-\overline{\alpha}_J)}\,d(r^{-\gamma})}{\int_{\|\boldsymbol{u}_J\|_p}^{\infty}(r^p-\|\boldsymbol{u}_J\|_p^p)^{\overline{\alpha}_I-1}r^{-p(\overline{\alpha}_I+\overline{\alpha}_J)+p}r^{-p(\overline{\alpha}-\overline{\alpha}_I-\overline{\alpha}_J)}\,d(r^{-\gamma})}\\ &=\frac{\int_{(\|\boldsymbol{u}_J\|_p^p+x^p)^{1/p}}^{\infty}(r^p-\|\boldsymbol{u}_J\|_p^p)^{\overline{\alpha}_I-1}r^{-p(\overline{\alpha}_I+\overline{\alpha}_J)+p}\,dF_\gamma(r)}{\int_{\|\boldsymbol{u}_J\|_p}^{\infty}(r^p-\|\boldsymbol{u}_J\|_p^p)^{\overline{\alpha}_I-1}r^{-p(\overline{\alpha}_I+\overline{\alpha}_J)+p}\,dF_\gamma(r)}\\ &=:\boldsymbol{P}\{\mathcal{R}_I>x\}.\end{aligned}$$

We note that

$$(r^p-\|\boldsymbol{u}_J\|_p^p)^{\overline{\alpha}_I-1}r^{-p\overline{\alpha}+p}\leq r^{p\overline{\alpha}_I-p}r^{-p\overline{\alpha}+p}\leq r^{-p\overline{\alpha}_J}\leq 1,\quad\forall r>0.$$

Applying now (44) and recalling that the random variable $R_{a_n}, n\geq 1$ is independent of $\boldsymbol{\mathcal{U}}$, we arrive at

$$c_n\Big(\boldsymbol{X}_I|\boldsymbol{X}_J=\boldsymbol{u}_{n,J}\Big)\quad\xrightarrow{d}\quad\mathcal{R}_I\boldsymbol{\mathcal{V}}_I,\quad n\to\infty.$$

In view of Theorem 5 we have

$$\mathcal{R}_I\boldsymbol{\mathcal{V}}_I\quad\stackrel{d}{=}\quad\boldsymbol{Y}_I|\boldsymbol{Y}_J=\boldsymbol{u}_J,$$

where $\boldsymbol{Y}\sim\mathcal{GSD}(k,p,\boldsymbol{\alpha},F_\gamma)$. This completes the proof.

□

Proof of Theorem 11: We shall show that statement i) implies that R is regularly varying with index $\gamma>0$. The rest of the proof follows along the lines of Theorem 3.1 of Hashorva (2006a).
Let $\boldsymbol{Z}:=(Z_1,\ldots,Z_k)^\top$ be a Kotz Type I LpGSD random vector in $\mathbb{R}^k$ with coefficient $\boldsymbol{\alpha}$ and $V>0$ be a random variable such that $V\sim Beta(\alpha_1,\overline{\alpha}-\alpha_1)$. Assume that $V,\boldsymbol{Z}$ and $\boldsymbol{X}$ are mutually independent and denote

$$R:=\|\boldsymbol{X}\|_p=\Big(\sum_{j=1}^k|X_i|^p\Big)^{1/p},\quad\widetilde{R}:=\|\boldsymbol{Z}\|_p=\Big(\sum_{j=1}^k|Z_i|^p\Big)^{1/p}.$$

By (13) $\widetilde{R}^{\,p}\sim Gamma(\overline{\alpha},1/p)$ Since X_1 is symmetric about 0, the assumptions that X_1 is regularly varying with index γ implies that $|X_1|^p$ is also regularly varying with index $\gamma/p>0$. $\widetilde{R}$ is independent of $|X_1|$, hence applying Lemma 17 (see

Appendix) we have that $(\widetilde{R}\,|X_1|)^p$ is also regular varying with positive index γ/p. In view of Corollary 3 we have

$$(\widetilde{R}\,|X_1|)^p \stackrel{d}{=} \widetilde{R}^{\,p}(R^pV) \stackrel{d}{=} (\widetilde{R}^{\,p}V)R^p \stackrel{d}{=} (|Z_1|R)^p,$$

consequently $(|Z_1|R)^p$ is regularly varying with index γ/p.
Applying once more (13) we have $|Z_1|^p \sim Gamma(\alpha_1, 1/p)$ with $|Z_1|^p$ being independent of R^p. Lemma 17 implies that R^p is regularly varying with the positive parameter γ/p. This concludes the proof.

□

Proof of Corollary 12: Denote $r(n) := F^{-1}(1-1/n), \forall n > 1$ with F^{-1} being the generalised inverse of the distribution function F. In view of Theorem 11, applying Proposition 0.8 (vii) of Resnick (1987), we have for $i \le k$

$$\lim_{n\to\infty} \frac{r(n)}{a_i(n)} \quad = \quad \left(\frac{\Gamma(\overline{\alpha}/p)\Gamma(\alpha_i+\gamma/p)}{2\Gamma(\alpha_i)\Gamma(\overline{\alpha}+\gamma/p)}\right)^{-1/\gamma} =: c_i.$$

Utilising the arguments presented in Theorem 11 we obtain for any $\boldsymbol{y} = (y_1, \ldots, y_k)^\top > \mathbf{0}$

$$\begin{aligned}
&\lim_{n\to\infty} n\Big[1 - \boldsymbol{P}\{Y_1/a_1(n) \le y_1, \ldots, Y_k/a_k(n) \le y_k\}\Big] \\
&= \lim_{n\to\infty} n\Big[1 - \boldsymbol{P}\Big\{\frac{Y_1}{r(n)}\frac{r(n)}{a_1(n)} \le y_1, \ldots, \frac{Y_k}{r(n)}\frac{r(n)}{a_k(n)} \le y_k\Big\}\Big] \\
&= \lim_{n\to\infty} n\Big[1 - \boldsymbol{P}\Big\{c_1\frac{Y_1}{r(n)} \le y_1, \ldots, c_k\frac{Y_k}{r(n)} \le y_k\Big\}\Big] \\
&= \gamma\int_0^\infty \boldsymbol{P}\{r\boldsymbol{c}A\boldsymbol{\mathcal{U}} \not\le \boldsymbol{y}\} r^{-\gamma-1}\,dr,
\end{aligned}$$

with $\boldsymbol{c} := (c_1, \ldots, c_k)^\top$ and $\boldsymbol{\mathcal{U}} \sim \mathcal{GSD}(k, p, \boldsymbol{\alpha})$. This completes the proof. □

8. Appendix

In this appendix several lemmas related to Dirichlet integrals, Gamma and Beta distributions are cited.

Lemma 13. *[Gupta and Song (1997), Lemma 1.1] Let $\boldsymbol{X}, \boldsymbol{Y}$ be two random vectors in $\mathbb{R}^k$ such that $\boldsymbol{X} \stackrel{d}{=} \boldsymbol{Y}$, and $f_i, 1 \le i \le d$, be measurable functions. We then have*

$$(f_1(\boldsymbol{X}), \ldots, f_d(\boldsymbol{X})) \stackrel{d}{=} (f_1(\boldsymbol{Y}), \ldots, f_d(\boldsymbol{Y})). \tag{45}$$

Lemma 14. *[Gupta and Song (1997), Lemma 2.3] Let f be a non-negative measurable function. For $\alpha_i > 0, i = 1, \ldots, k$, we have:*

$$\int_{[0,\infty)^k} f(\sum_{i=1}^k x_i)\prod_{i=1}^k x_i^{\alpha_i-1}\,dx_1\cdots dx_k \quad = \quad \frac{\prod_{i=1}^k \Gamma(\alpha_i)}{\Gamma(\overline{\alpha})}\int_0^\infty f(x)x^{\overline{\alpha}-1}\,dx, \tag{46}$$

provided one of the integrals exist.

The next lemma is a minor generalisation of Lemma 2.3 of Gupta and Song (1997).

Lemma 15. *Let f be a non-negative measurable function. We then have for any $p_i > 0$ and $\alpha_i > 0, i = 1, \ldots, k$,*

$$\int_{[0,\infty)^k} f(\sum_{i=1}^k x_i^{p_i}) \prod_{i=1}^k x_i^{\alpha_i - 1}\, dx_1 \cdots dx_k = \frac{\prod_{i=1}^k \Gamma(\alpha_i/p_i)}{\Gamma(\overline{\alpha}/p_i) \prod_{i=1}^k p_i} \int_0^\infty f(x) x^{\overline{\alpha}/p_i - 1}\, dx \tag{47}$$

and

$$\int_{\mathbb{R}^k} f(\sum_{i=1}^k |x_i|^{p_i}) \prod_{i=1}^k |x|_i^{\alpha_i - 1}\, dx_1 \cdots dx_k = \frac{2^k \prod_{i=1}^k \Gamma(\alpha_i/p_i)}{\Gamma(\overline{\alpha}/p_i) \prod_{i=1}^k p_i} \int_0^\infty f(x) x^{\overline{\alpha}/p_i - 1}\, dx \tag{48}$$

provided one of the integrals exist.

Proof. Assume that the integral

$$I := \int_{[0,\infty)^k} f(\sum_{i=1}^k x_i^{p_i}) \prod_{i=1}^k x_i^{\alpha_i - 1}\, dx_1 \cdots dx_k$$

is finite. Changing the variables $y_i := x_i^{p_i}, i \le k$, and using Lemma 14 we obtain

$$\begin{aligned} I &= \frac{1}{\prod_{i=1}^k p_i} \int_{[0,\infty)^k} f(\sum_{i=1}^k y_i) \prod_{i=1}^k y_i^{\alpha_i/p_i - 1}\, dy_1 \cdots dy_k \\ &= \frac{\prod_{i=1}^k \Gamma(\alpha_i/p_i)}{\Gamma(\overline{\alpha}/p_i) \prod_{i=1}^k p_i} \int_0^\infty f(x) x^{\overline{\alpha}/p_i - 1}\, dx. \end{aligned}$$

Now the equation (48) follows easily.

□

Theorem 16. *Let Y be a random variable with the distribution function H which has the upper endpoint $\omega \in (0, \infty]$ and $H(0) = 0$. Let $Z_{a,b}$ be a Beta distributed random variable with positive parameters a, b being independent of Y, and $\tau > 0$ be a fixed constant.*
i) If $H \in MDA(\Lambda)$ with a positive scaling function w we have as $u \uparrow \omega$

$$\boldsymbol{P}\{Y[1 - Z_{a,b}]^{1/\tau} > u\} = (1 + o(1)) \frac{\Gamma(a+b)}{\Gamma(b)} \left(\frac{\tau}{uw(u)} \right)^a [1 - H(u)]. \tag{49}$$

ii) If $H \in MDA(\Phi_\alpha), \alpha > 0$, then $\omega = \infty$ and

$$\boldsymbol{P}\{Y[1 - Z_{a,b}]^{1/\tau} > u\} = (1 + o(1)) \frac{\Gamma(a+b)\Gamma(b + \alpha/\tau)}{\Gamma(b)\Gamma(a + b + \alpha/\tau)} [1 - H(u)] \tag{50}$$

holds as $u \to \infty$
iii) If $H \in MDA(\Psi_\alpha), \alpha > 0$ and $\omega = 1$, we then have

$$\begin{aligned} &\boldsymbol{P}\{Y[1 - Z_{a,b}]^{1/\tau} > u\} \\ &\quad = (1 + o(1)) \frac{\Gamma(\alpha+1)\Gamma(a+b)}{\Gamma(b)\Gamma(\alpha + a + 1)} (\tau(1-u))^a [1 - H(u)], \quad u \uparrow 1. \end{aligned} \tag{51}$$

Proof. The proof for the case $\tau = 2$ is given in Theorem 12.3.1, 12.3.2, 12.3.3 of Berman (1992). The general case $\tau > 0$ is shown in Theorem 6.2 of Hashorva (2006d). □

Lemma 17. *Let X, Y be two independent positive random variables with $Y^p \sim Gamma(a,\lambda), a,\lambda>0, p>0$. If X is regularly varying with index $\gamma \geq 0$, we then have*

$$\lim_{u\to\infty} \frac{\boldsymbol{P}\{XY>u\}}{\boldsymbol{P}\{Y>u\}} = \frac{\Gamma(a+\gamma/p)}{\lambda^{\gamma/p}\Gamma(a)} \in (0,\infty). \tag{52}$$

Conversely, if the product XY is regularly varying with index $\gamma \geq 0$, then X is also regularly varying with index γ and furthermore (52) *is valid.*

Proof. The proof can be found in Lemma 6.1 of Hashorva (2006d) where the case $\gamma>0$ is considered. We sketch it below. If X is regularly varying with the positive index $\gamma \geq 0$, then (52) follows by Breiman's Lemma (see for some deep related results Denis and Zwart (2005)).
Suppose for simplicity that $p=1, \lambda=1$. For any $t>0$ we may write by the independence of X and Y

$$\begin{aligned}\boldsymbol{P}\{XY>t\} &= \frac{t^a}{\Gamma(a)}\int_0^\infty \boldsymbol{P}\{XY>t|Y=tx\}\exp(-tx)x^{a-1}\,ds\\ &= t^a\int_0^\infty \exp(-tv)\,dG(v),\end{aligned}$$

where

$$G(s) := \frac{1}{\Gamma(a)}\int_0^s \boldsymbol{P}\{X>1/x\}x^{a-1}\,dx, \quad s>0.$$

The assumption XY is regularly varying with index $\gamma \geq 0$ means

$$\int_0^\infty \exp(-tv)\,dG(v) = t^{-a-\gamma}L(1/t), \quad t\to\infty, \tag{53}$$

with $L(x)$ such that $\lim_{t\to 0} L(Kt)/L(t)=1, \forall K>0$. In view of Karamata's Tauberian Theorem (Feller (1966), Resnick (1987)) (53) is equivalent with

$$G(t) = \frac{1}{\Gamma(a+\gamma+1)}t^{a+\gamma}L(t), \quad t\to 0,$$

or equivalently

$$G(1/t) = \frac{1}{\Gamma(a+\gamma+1)}t^{-a-\gamma}L(1/t), \quad t\to\infty.$$

Consequently as $t\to\infty$

$$\int_0^{1/t} \boldsymbol{P}\{X>1/x\}x^{a-1}\,dx = \frac{\Gamma(a)}{\Gamma(a+\gamma+1)}t^{-a-\gamma}L(1/t).$$

Since $\boldsymbol{P}\{X>x\}x^{-a-1}, x>0$ decreases monotonically in x for any $a>0$ we get applying the Monotone Density Theorem (Resnick (1987))

$$\boldsymbol{P}\{X>t\}t^{-a-1} = \frac{(a+\gamma+1)\Gamma(a)}{\Gamma(a+\gamma+1)}t^{-a-\gamma-1}L(1/t), \quad t\to\infty,$$

thus the proof follows. □

References

[1] Anderson, T.W., and Fang, K.T. (1990) On the theory of multivariate elliptically contoured distributions and their applications. In *Statistical Inference in Elliptically Contoured and Related Distributions*, K.T. Fang and T.W. Anderson, eds, Allerton Press, New York, pp. 1–23.

[2] Basrak, B., Davis, R.A., and Mikosch, T. (2002) A characterization of multivariate regular variation. *Ann. Appl. Probab.* **12**,3, 908–920.

[3] Berman, M.S. (1982) Sojourns and extremes of stationary processes. *Ann. Probab.* **10**, 1–46.

[4] Berman, M.S. (1983) Sojourns and extremes of Fourier sums and series with random coefficients. *Stoch. Proc. Appl.* **15**, 213–238.

[5] Berman, M.S. (1992) *Sojourns and Extremes of Stochastic Processes.* Wadsworth & Brooks/ Cole, Boston.

[6] Cambanis, S., Huang, S., and Simons, G. (1981) On the theory of elliptically contuored distributions. *J. Multivariate Anal.* **11**, 368–385.

[7] de Haan, L. (1970) *On Regular Variation and its Applications to the Weak Convergence of Sample Extremes.* Mathematisch Centrum Amsterdam, The Netherlands.

[8] Denis, D., and Zwart, B. (2005) On a theorem of Breiman and a class of random difference equations. Preprint.

[9] Dirichlét, P.G.L. (1839) Sur un nouvelle méthode pour le determination des intégrales multiples. *Liouville Journal des Mathematiques*, Ser I, 4, 164–168.

[10] Falk, M., Hüsler, J., and Reiss R.-D. (2004) *Laws of Small Numbers: Extremes and Rare Events.* DMV Seminar **23**, 2-nd edition, Birkhäuser, Basel.

[11] Fang, K.-T., and Fang, Bi-Qi. (1990) Generalised symmetrised Dirichlet distributions. In *Statistical Inference in Elliptically Contoured and Related Distributions*, K.T. Fang and T.W. Anderson, eds, Allerton Press, New York, pp. 127–136.

[12] Fang, K.-T., Kotz, S., and Ng, K.-W. (1990) *Symmetric Multivariate and Related Distributions.* Chapman and Hall, London, United Kingdom.

[13] Fang, K., and Zhang, Y. (1990) *Generalized Multivariate Analysis.* Springer, Berlin, Heidelberg, New York.

[14] Gupta, A.K., and Song, D. (1997) L_p-norm spherical distributions, *J. Statist. Plann. Inference*, 60, 241–260.

[15] Gupta, A. K., and Varga, T. (1993) *Elliptically Contoured Models in Statistics.* Kluwer, Dordrecht.

[16] Hashorva, E. (2005a) Extremes of asymptotically spherical and elliptical random vectors. *Insurance: Mathematics and Economics*, **36**, 3, 285–302.

[17] Hashorva, E. (2006a) On the regular variation of elliptical random vectors. *Stat. Probab. Letter.* **76**, 14, 1427–1434.

[18] Hashorva, E. (2006b) Gaussian approximation of conditional elliptical random vectors. *Stochastic Models.* **22**, 441-457.

[19] Hashorva, E. (2006c) On the multivariate extremes and asymptotic dependence of elliptical random vectors. Preprint.

[20] Hashorva, E. (2006d) Extremes of L_p-norm asymptotically spherical distributions. Preprint.

[21] Hashorva, E. (2007) Conditional limiting distribution of type III elliptical random vectors. *J. Multivariate Anal.* **98**, 282–194.

[22] Kallenberg, O. (1997) *Foundations of Modern Probability.* New York, Springer.

[23] Kano, Y. (1994) Consistency property of elliptical probability density functions. *J. Multivariate Anal.* 51, 139–147.

[24] Kotz, S. (1975) Multivariate distributions at a cross-road. In: *Statistical Distributions in Scientific Work* 1, G.P. Patil, S. Kotz, and J.K. Ord eds, D. Riedel, Dortrecht, 240–247.

[25] Kotz, S., Balakrishnan, N., and Johnson, N.L. (2000) *Continuous Multivariate Distributions.* Second Edition, Wiley, New York.

[26] Kotz, S., and Ostrovskii, I.V. (1994) Characteristic functions of a class of elliptical distributions. *J. Multivariate Anal.* **49**, (1), 164–178.

[27] Kotz, S., and K.W. Ng (1995) Some new classes of multivariate Louville distribution. In *Proc. of 50th Session of the International Statistical Institute* (Section CP 46), Beijing, China.

[28] Kotz, S., and Nadarajah, S. (2005) *Extreme Value Distributions, Theory and Applications.* Imperial College Press, London, United Kingdom. (Second Printing).

[29] Leadbetter, M.R., Lindgren, G., and Rootzén, H. (1983) *Extremes and related properties of random sequences and processes.* Springer-Verlag, New York.
[30] Reiss, R-D. (1989) *Approximate Distributions of Order Statistics: With Applications to Non-parametric Statistics.* Springer, New York.
[31] Resnick, S.I. (1987) *Extreme Values, Regular Variation and Point Processes.* Springer, New York.
[32] Szabłowski, P.L. (1990) Expansions of $E(X|Y+\epsilon X)$ and their applications to the analysis of elliptically contoured measures. (English) *Comput. Math. Appl.* 19, No.5, 75–83.
[33] Szabłowski, P.L. (1998) Uniform distributions on spheres in finite dimensional L_α and their generalizations. *J. Multivariate Anal.* **64**, 103–117.

University of Bern, Institute of Mathematics Statistics and Actuarial Sciences,, Sidlerstrasse 5, CH-3012 Bern, Switzerland, and, Allianz Suisse Insurance Company, Bern, Switzerland, Laupenstrasse 27, CH-3001 Bern, Switzerland
E-mail address: enkelejd.hashorva@stat.unibe.ch

The George Washington University, School of Engineering & Applied Science, 1776 G Street NW, Suite 110, Washington, D.C. 20052, USA
E-mail address: kotz@gwu.edu

University of Kent, Institute of Mathematics Statistics and Actuarial Sciences, CT2 7NF United Kingdom
E-mail address: a.kume@kent.ac.uk

ALBANIAN JOURNAL
OF MATHEMATICS
Volume 1, Number 1, Pages 57–62
ISSN 1930-1235: (electronic version)

ABSOLUTELY SUMMING OPERATORS IN $m_1(l_1)$

NAIM L. BRAHA

Abstract. A scalar sequence (a_i) is said to be a p-multiplier of a Banach space X, if it satisfies the following condition:

$$m_p(X) = \left\{ a = (a_i) : \sum_i \|a_i x_i\|^p < \infty, \forall (x_i) \in {l_w}^p(X) \right\}.$$

In this paper we will prove the following: every bounded linear operator from Banach space l_1 into $m_1(l_1)$, is an absolutely summing operator.

1. Introduction

The theory of absolutely summing operators has as a starting point from the famous resume of Grothendick [5], in which it was proved that every bounded linear operator from Banach space l_1 into l_2 is an absolutely summing operator. In later works by Pietsch necessary and sufficient conditions are given under which an operator is an absolutely summing operator; see [4], [6]. In this context the absolutely summing operators were studied in the sequence spaces and function spaces by several authors; see [4] for further references.The sequence space $m_1(l_1)$ was defined by the authors S. Aywa and J. H. Fourie in [2]. In this paper we prove that every bounded linear operator from Banach space l_1 into space $m_1(l_1)$, is an absolutely summing operator. We also prove that every bounded linear operator from Banach space l_1 into $m_1(X)$ is an absolutely summing operator, after taking in consideration the definition of the 1-colacunary sequences given in [1] and their properties. Also we give a result which characterizes the absolutely summing operators from space $m_1(X)$ into l_2, in case where X contains a basis which satisfies the 1-colacunarity.

2. Preliminaries

In the first part of the paper we will prove that every bounded linear operator from Banach space l_1 into $m_1(l_1)$, is an absolutely summing operator. In the second part we will prove the following: If X contains a basic and 1-colacunary vector sequence (x_i), then every bounded linear operator from l_1 into $m_1(X)$ is an absolutely summing operator. In the sequel we will briefly describe the notation and definitions which are used throughout the paper.

Let Λ denote the vector space of scalar sequences (a_i), where (a_i) are from $\mathbb{R}$ or $\mathbb{C}$, i.e.,

$$\Lambda = \{a = (a_i) : a_i \in \mathbb{R} \quad \text{or} \quad a_i \in \mathbb{C}\}.$$

Received by the editors August 8, 2006 and, in revised form, February 15, 2007.
2000 *Mathematics Subject Classification.* Primary 47B10,47B37; Secondary 47B38.
Key words and phrases. scalar sequences, absolutely summing operators.

The space $m_p(X)$, is defined as follows

$$m_p(X) = \left\{ a = (a_i) \in \Lambda : \sum_i ||a_i x_i||^p < \infty, \forall (x_i) \in {l_w}^p(X) \right\}, \tag{1}$$

and is a Banach space under the norm

$$||(a_i)||_{p,p} = \sup_{\epsilon_p((x_i))\leq 1} \left(\sum_{n\in\mathbb{N}} |a_n|^p ||x_n||^p \right)^{\frac{1}{p}},$$

where $\epsilon_p((x_i)) = \sup_{||a||\leq 1} ||a(x_i)||_p$, $a \in X^*$ (see [2]).

By ${l_w}^p(X)$ we will denote the Banach space

$${l_w}^p(X) = \left\{ x = (x_i) \in X : \left(\sum_i |x^*(x_i)|^p \right)^{\frac{1}{p}} < \infty, x^* \in X^* \right\}.$$

For the class of the scalar sequences $m_p(X)$, the following inclusion holds

$$l_p \subseteq m_p(X) \subset l_\infty,$$

for any $1 \leq p \leq \infty$.

Definition 1. *Let X be a Banach space. A sequence $(x_n)_{n\in\mathbb{N}}$ in X is p-colacunary if there is a $\delta > 0$ such that*

$$\left\| \sum_{i\leq n} a_i x_i \right\| \geq \delta \left(\sum_{i\leq n} |a_i|^p \right)^{\frac{1}{p}},$$

for any sequence of scalars $a_0, a_1, \cdots, a_n$.

The following theorem is proved in [5],[6].

Theorem 1. *Every bounded linear operator defined from Banach space l_1 into space l_2, is an absolutely summing operator.*

The next result is known as the "Ideal property of p-summing operators"; see [4] for details.

Theorem 2. *Let $1 \leq p < \infty$ and $v \in \Pi_p(X, Y)$. Then the composition of v with any bounded linear operator is p-summing.*

3. Results

Theorem 3. *Every bounded linear operator from Banach space l_1 into Banach space $m_1(l_1)$, is an absolutely summing operator.*

Proof. From the facts mentioned above, in order to prove the Theorem, it is enough to prove the fact that the Banach space $m_1(l_1)$, is a subspace of the space l_2 and the norm in $m_1(l_1)$ is equivalent with the standard norm given in l_2. Let us consider that $a = (a_i)$, is any scalar sequence from space $m_1(l_1)$,

$$m_1(l_1) = \left\{ a = (a_i) \in \Lambda : \sum_i ||a_i x_i|| < \infty, \forall (x_i) \in {l_w}^1(l_1) \right\}. \tag{2}$$

Let us denote by (e_i) the standard unit vector basis in l_1, and let us define the operator A from l_1 into l_1, by the following relation:

$$A : x = \sum_i b_i e_i \to \sum_i a_i b_i e_i,$$

for any sequence $(a_i) \in m_1(l_1)$. The above operator is well defined, because $(a_i b_i) \in l_1$. Indeed, from the above it was shown that $m_1(l_1) \subset l_\infty$ and from this follows that the following relation is true,

$$\sum_i |a_i \cdot b_i| \leq \sup_i |a_i| \sum_i |b_i| < \infty.$$

We have

$$\begin{aligned} ||Ax|| &= \left\| A\left(\sum_i b_i \cdot e_i\right)\right\| = \left\|\left(\sum_i b_i \cdot a_i \cdot e_i\right)\right\| \\ &= \sum_i |b_i \cdot a_i| \leq \sup_i |a_i| \cdot \sum_i |b_i| \\ &= \sup_i |a_i| \cdot \left\|\sum_i b_i \cdot e_i\right\| = \sup_i |a_i| \cdot ||x||. \end{aligned}$$

Thus, it follows that the operator A is a bounded linear operator. Hence, we have

$$\sum_{i\in\mathbb{N}} ||A(a_i x_i)|| \leq S \sum_{i\in\mathbb{N}} ||a_i x_i|| < \infty, \tag{3}$$

where $S = \sup_i |a_i|$ for any $(a_i) \in m_1(l_1)$ and $(x_i) \in l_1$. Without loss of generality we can assume that the sequence of vectors (x_i) is normalized. Taking in consideration the relation (3) we have:

$$\sum_{i\in\mathbb{N}} ||A(a_i x_i)|| = \sum_{i\in\mathbb{N}} |a_i| \cdot ||A(x_i)|| = \sum_{i\in\mathbb{N}} |a_i| \cdot ||a_i x_i|| = \sum_{i\in\mathbb{N}} {a_i}^2 < \infty.$$

The last relation proves that $(a_i) \in l_2$. Next, we aim to prove that the norm $||(a_i)||_{1,1}$ is equivalent with $||(a_i)||_{l_2}$. Let $(a_i) \in m_1(l_1)$. Then

$$\begin{aligned} ||(a_i)||_{1,1} &= \sup_{\epsilon_1((x_i))\leq 1} \sum_{n\in\mathbb{N}} |a_n| \cdot ||x_n|| \leq \sup_n |a_n| \sup_{\epsilon_1((x_i))\leq 1} \sum_{n\in\mathbb{N}} ||x_n|| \leq \\ &\left(\sum_{n\in\mathbb{N}} |a_n|^2\right)^{\frac{1}{2}} \cdot N = N \cdot ||(a_n)||_{l_2}, \end{aligned} \tag{4}$$

$\forall (x_n) \in {l_w}^1(l_1)$. From Schur's l_1–theorem (see [4] for details) it follows that

$$N = \sup_{\epsilon_1((x_i))\leq 1} \sum_{n\in\mathbb{N}} ||x_n|| < \infty.$$

Take $(a_i) \in l_2$ and consider that $\sum_{i=1}^n {a_i}^2 = 1$. Then, from Dvoretzky-Rogers theorem it follows that there exists an unconditional sequence $(y_n) \in l_1$, such that $||y_n||_{l_1} = |a_n|$; see [6] for details. From unconditionality of (y_n), we have that $\sum_{i\in\mathbb{N}} \theta_i y_i < \infty$, for any sequence of signs (θ_i) (see [6, Prop.1, c.1]), respectively $\sum_{i\in\mathbb{N}} |y_i| < \infty$, from which follows that it converges $\sum_{i\in\mathbb{N}} |y^*(y_i)| < \infty$ and

$\sum_{i\in\mathbb{N}} |a_i|\cdot||y_i|| = \sum_{i\in\mathbb{N}} |a_i|^2 < \infty$. Which means that the relation

$$\sup_{\epsilon_1(y_i)\leq 1} \sum_{n\in\mathbb{N}} |a_n|\cdot||y_n||,$$

defines a norm $||(a_i)||_{1,1}$ on $m_1(l_1)$. Now we have the estimation,

$$||(a_n)||_{l_2} = \left(\sum_{n\in\mathbb{N}} |a_n|^2\right)^{\frac{1}{2}} = \left(\sum_{n\in\mathbb{N}} |a_n|\cdot||y_n||\right)^{\frac{1}{2}} \leq$$

$$\left(\sup_{\epsilon_1((y_i))\leq 1} \sum_{n\in\mathbb{N}} |a_n|||y_n||\right)^{\frac{1}{2}} \leq ||(a_n)||_{1,1}. \tag{5}$$

From relations (4) and (5) it follows that the norms $||(a_i)||_{l_2}$ and $||(a_i)||_{1,1}$ are equivalent. This completes the proof.

□

Proposition 4. *Let $(x_n)_{n\in\mathbb{N}}$ be a basic and 1-colacunary sequence of vectors in a Banach space X.Then every bounded linear operator T from l_1, into $m_1(X)$, is an absolutely summing operator.*

Proof. Let (x_n), be a 1-colacunary sequence of vectors in Banach space X, then there follows the following relation

$$\delta\cdot\sum_{i\leq n}|a_i| \leq \left\|\sum_{i\leq n} a_i x_i\right\| \leq \sum_{i\leq n}|a_i|,$$

which means that sequence of vectors (x_n) is equivalent with standard unit vector basis of l_1. The rest of the proof is similar to that of Theorem 3.

□

Theorem 5. *Let $(x_n)_{n\in\mathbb{N}}$ be a basic and 1-colacunary sequence of vectors in a Banach space X. Then every bounded linear operator T from $m_1(X)$, into l_2, is an absolutely summing operator.*

Proof. Let us denote by (f_i) the basic sequence in Banach space $m_1(X)$, and let (x_i) be a basic and 1-colacunary sequence of vectors in X. Then,

$$\left\|\sum_{i\leq n} a_i x_i\right\| \geq \delta\sum_{i\leq n}|a_i|,$$

for any finite sequence of scalars $a_0, a_1, \cdots, a_n$. From this relation it follows that

$$\delta\cdot\sum_{i\leq n}|a_i| \leq \left\|\sum_{i\leq n} a_i x_i\right\| \leq \sum_{i\leq n}|a_i|, \tag{6}$$

for any sequence of scalars $a_0, a_1, \cdots, a_n$. Let A be an operator defined from l_1 into $m_1(X)$, by the relation

$$A : x = \sum_i a_i e_i \to \sum_i a_i f_i,$$

where (e_i) is the standard unit vector basis in l_1. The operator A is well defined. Next, we prove that A is bounded from the upper side, lower side, and bijective, from which it follows that it has bounded inverse A^{-1}; see [8]. We have

$$||Ax|| = \left\| A\left(\sum_i a_i e_i\right)\right\| = \left\|\sum_i a_i f_i\right\| = \|(a_i)\|_{m_1(X)}$$
$$= \sup_{\epsilon_1((y_i))\le 1}\left(\sum_n |a_n|^1 \cdot ||y_n||^1\right) \le$$
$$\sup_{\epsilon_1((y_i))\le 1} \sup_n ||y_n|| \cdot \sum_n |a_n| = M \cdot \left\|\sum_i a_i e_i\right\|_{l_1} = M \cdot ||x||,$$

where $(y_i) \in {l_w}^1(X)$ and $M = \sup_{\epsilon_1((y_i))\le 1} \sup_n ||y_n||$. In the similar way we can prove the lower bound of $||Ax||$. In the following, it is enough to prove that A is onto (because injectivity follows from the definition). Let $y = (c_i) \in m_1(X)$ be any element from that space, then it is enough to prove that there follows

$$\sum_i |c_i| < \infty.$$

Relation (6) is true for any scalar sequence (a_i), so it remains true if we are using the scalar sequence (c_i), instead (a_i) i.e., the following relation is valid

$$\sum_i |c_i| \le \frac{1}{\delta} \cdot \left\|\sum_i c_i x_i\right\| \le \frac{1}{\delta} \cdot \sum_i ||c_i x_i|| < \infty.$$

This proved that A is a bijective operator with bounded inverse. The following diagram is commutative

$$\begin{array}{ccc} l_1 & \xrightarrow{A} & m_1(X) \\ & \searrow^{C} & \downarrow{\scriptstyle T} \\ & & l_2 \end{array}$$

Let C denote the operator which is the composition of the operators A and T, i.e.,

$$C = T \cdot A. \tag{7}$$

The operator C is a bounded linear operator from Banach space l_1 into space l_2, so it is absolutely summing operator between them (Theorem 1). Then from relation (7) we will have that $C \cdot A^{-1} = T$. From Theorem 2, follows that the operator T is also an absolutely summing operator.

□

The proof of the following Proposition is similar to that of Proposition 10 in [3].

Proposition 6. *Let $(x_n)_{n\in\mathbb{N}}$ be a basic and 1-colacunary sequence of vectors in X. Then every infinite dimensional subspace Y of $m_1(X)$ is isomorphic to $m_1(X)$ and complemented in $m_1(X)$. Hence, the Banach space $m_1(X)$ is a Prime space.*

Proof. Let H be an operator defined from the Banach space $m_1(X)$ into the space l_1 by the relation

$$H : x = \sum_i a_i f_i \to \sum_i a_i e_i,$$

where (f_i) and (e_i), are basic sequences in $m_1(X)$, l_1 respectively. This operator is invertible (exactly as operator A in Theorem 4). Let Y be any infinite dimensional subspace of $m_1(X)$. Let us denote by $Y_1 = H(Y)$, the subspace of l_1. From the decomposition method of Pelczynski it follows that

$$l_1 = Y_1 \oplus B$$

for some Banach space B; see [7]. Let $x \in m_1(X)$. Then $H(x) = y \in l_1$ and y has unique representation

$$y = a + b \tag{8}$$

for suitable $a \in Y_1$ and $b \in B$. From this there is a $a_1 \in Y, H(a_1) = a$,

$$y = H(a_1) + b \Rightarrow H^{-1}(y) = H^{-1}(H(a_1)) + H^{-1}(b) \Rightarrow$$

$$x = a_1 + H^{-1}(b) \tag{9}$$

and the last representation of x is unique. If we use another representation of x we will have $x = a_1^{'} + H^{-1}(b^{'})$, then $H(x) = H(a_1^{'}) + b^{'} \Rightarrow$

$$y = H(a_1^{'}) + b^{'}. \tag{10}$$

But relation (10) is in contradiction with relation (8). So every $x \in m_1(X)$ has unique representation through space Y, and we can use the notation

$$m_1(X) = Y \oplus C$$

for some Banach space C, with Y isomorphic to $m_1(X)$. Thus, $H(Y) = Y_1$ is isomorphic to l_1. Let us denote by B that isomorphism between them. Then $B(l_1) = BH(m_1(X)) = Y_1 \Rightarrow BH(m_1(X)) = H(Y)$ and from this follows that $H^{-1} \cdot B \cdot H$ is isomorphism between spaces $m_1(X)$ and Y. This completes the proof.

□

Corollary 7. *The space $m_1(l_1)$ is a Prime space.*

References

[1] D. J. Aldous and D. H. Fremlin, Colacunary sequences in L-spaces, Studia Math. T. 71. (1982), 297-304.

[2] S. Aywa and J. H. Fourie, On summing Multipliers and Aplications, J.Math.Anal.Appl. 253, (2001), 166-186.

[3] N. L. Braha, Characterization of the absolutely summing operators in a Banach space using μ -approximate l_1 sequences, Matematiche, Vol. LX (2005)-Fasc. I, 121-128.

[4] J. Diestel, H. Jarchow and A. Tonge, Absolutely summing operators, Cambridge Univ.Press, 1995.

[5] A. Grothendieck, Resume de la theorie metrique des produits tensoriel topologiques, Bol.Soc.Mat. Sao Paulo 8 (1953/56), 1-79.

[6] J. Lindenstrauss and L. Tzafriri, Classical Banach spaces, Part I, Springer-Verlag, Berlin, Heidelberg, New York 1977.

[7] A. Pelczynski, Projections in certain Banach spaces, Studia Math.19 (1960), 209-228.

[8] W. Rudin, Functional analysis, McGraw-Hill, 1973.

Department of Mathematics and Computer Sciences,, Avenue "Mother Theresa " 5, Prishtinë, 10 000,, University of Prishtina, Prishtina, Kosova.

ALBANIAN JOURNAL
OF MATHEMATICS
Volume 1, Number 2, Pages 67–85
ISSN 1930-1235: (electronic version)

GROUP REPRESENTATIONS ON RIEMANN-ROCH SPACES OF SOME HURWITZ CURVES

DAVID JOYNER, AMY KSIR, AND ROGER VOGELER

(Communicated by T. Shaska)

ABSTRACT. Let $q > 1$ denote an integer relatively prime to $2, 3, 7$ and for which $G = PSL(2, q)$ is a Hurwitz group for a smooth projective curve X defined over $\mathbb{C}$. We compute the G-module structure of the Riemann-Roch space $L(D)$, where D is an invariant divisor on X of positive degree. This depends on a computation of the ramification module, which we give explicitly. In particular, we obtain the decomposition of $H^1(X, \mathbb{C})$ as a G-module.

1. INTRODUCTION

Let X be a smooth projective curve over an algebraically closed field k, and let $k(X)$ denote the function field of X (the field of rational functions on X). If D is any divisor on X then the Riemann-Roch space $L(D)$ is a finite dimensional k-vector space given by

$$L(D) = L_X(D) = \{f \in k(X)^\times \mid \mathrm{div}(f) + D \geq 0\} \cup \{0\},$$

where $\mathrm{div}(f)$ denotes the (principal) divisor of the function $f \in k(X)$. If G is a finite group of automorphisms of X, then G has a natural action on $k(X)$, and on the group $\mathrm{Div}(X)$ of divisors on X. If D is a G-invariant divisor, then G also acts on the vector space $L(D)$, making it into a $k[G]$-module.

The problem of finding the $k[G]$-module structure of $L(D)$ was first considered in the case where $k = \mathbb{C}$ and D is canonical, i.e. $L(D)$ is the space of holomorphic differentials on X. This problem was solved by Hurwitz for G cyclic, and then by Chevalley and Weil for general G. More generally, the problem has been solved by work of Ellingsrud and Lønsted [EL], Kani [K], Nakajima [N], and Borne [B]. This has resulted in the following equivariant Riemann-Roch formula for the class of $L(D)$ (denoted by square brackets) in the Grothendieck group $R_k(G)$, in the case where D is non-special:

$$[L(D)] = (1 - g_{X/G})[k[G]] + [\mathrm{deg}_{eq}(D)] - [\tilde{\Gamma}_G]. \tag{1}$$

Here $g_{X/G}$ is the genus of X/G, $\mathrm{deg}_{eq}(D)$ is the equivariant degree of D, and $\tilde{\Gamma}_G$ is the (reduced) ramification module (this notation will be defined in sections 4.1 and 4.2).

Received by the editors March 21, 2007.
2000 *Mathematics Subject Classification.* 14Q05, 14H37, 14F40, 14H45.
Key words and phrases. Hurwitz curves, representations.

Explicitly computing the $k[G]$-module structure of $L(D)$ in specific cases is of interest currently due to advances in the theory of algebraic-geometric codes. Permutation decoding algorithms use this information to increase their efficiency.

In this paper, we consider the case where X is a Hurwitz curve with automorphism group $G = PSL(2,q)$ for some prime power q, over $k = \mathbb{C}$. Using the equivariant Riemann-Roch formula above (1) and the representation theory of $PSL(2,q)$, we compute explicitly the $\mathbb{C}[G]$-module structure of $L(D)$ for a general invariant effective divisor D. In the case where D is a canonical divisor, this yields an explicit computation for the $\mathbb{C}[G]$-module structure of $H^1(X,\mathbb{C})$.

We are also interested in rationality questions. We find that $\tilde{\Gamma}_G$ has a $\mathbb{Q}[G]$-module structure, and therefore may be computed more simply (see Joyner and Ksir [JK1]), as follows:

$$\tilde{\Gamma}_G = \bigoplus_{\pi \in G^*} \left[\sum_{\ell=1}^{L} (\dim \pi - \dim (\pi^{H_\ell})) \frac{R_\ell}{2} \right] \pi. \tag{2}$$

The sum is over all conjugacy classes of cyclic subgroups of G, H_ℓ is a representative cyclic subgroup, π^{H_ℓ} indicates the fixed part of π under the action of H_ℓ, and R_ℓ denotes the number of branch points in Y over which the decomposition group is conjugate to H_ℓ. For some but not all divisors D, $L(D)$ has a $\mathbb{Q}[G]$-module structure, and may also be computed more simply.

The organization of this paper is as follows. In section 2, we recall some facts about Hurwitz curves and Hurwitz groups. In section 3, we review the representation theory of $PSL(2,q)$, and compute the induced characters necessary for the following section. Our main results are in section 4, where we compute the ramification module, the equivariant degree for any invariant divisor D, and thus the structure of $L(D)$. At the end of section 4 we compute the $\mathbb{C}[G]$-module structure of $H^1(X,\mathbb{C})$. In section 5, we discuss rationality questions, using the results of [JK1] to give more streamlined formulas for the ramification module, and in some cases for $L(D)$.

2. Hurwitz curves

The automorphism group G of a smooth projective curve of genus $g > 1$ over an algebraically closed field k of characteristic zero satisfies the *Hurwitz bound*

$$|G| \leq 84 \cdot (g-1).$$

A curve which attains this bound is called a *Hurwitz curve* and its automorphism group is called a *Hurwitz group*.

2.1. **Classification.** The number of distinct Hurwitz groups is infinite, and to each one corresponds a finite number of Hurwitz curves. Nevertheless, these curves are quite rare; in particular, the Hurwitz genus values are known to form a rather sparse set of positive integers (see Larsen [L]).

Hurwitz groups are precisely those groups which occur as non-trivial finite homomorphic images of the 2,3,7-triangle group

$$\Delta = \langle a, b : a^2 = b^3 = (ab)^7 = 1 \rangle.$$

This is most naturally viewed as the group of orientation-preserving symmetries of the tiling of the hyperbolic plane $\mathbf{H}$ generated by reflections in the sides of a fundamental triangle having angles $\pi/2$, $\pi/3$, and $\pi/7$. Each proper normal

finite-index subgroup $K \triangleleft \Delta$ corresponds to a Hurwitz group $G = \Delta/K$. The associated Hurwitz curve now appears (with $k = \mathbb{C}$) as a compact hyperbolic surface $\mathbf{H}/K$ regularly tiled by a finite number of copies of the fundamental triangle. G is the group of orientation-preserving symmetries of this tiling, with fundamental domain consisting of one fundamental triangle plus one reflected triangle. (From this perspective, the Hurwitz bound simply says that there is no smaller polygon which gives a regular tiling of $\mathbf{H}$.)

We note that Δ has only a small number of torsion elements (up to conjugacy). These are the non-trivial powers of a, b, and ab. Each acts as a rotation of order 2, 3, or 7, and has as its fixed point one vertex of (some copy of) the fundamental triangle. Clearly no other point of the tiling can occur as a fixed point; this is true both for the tiling of $\mathbf{H}$ and the induced tilings on the quotient surfaces. In other words, all points *other* than the tiling vertices have trivial stabilizer.

It follows easily from the above presentation for Δ that a group is Hurwitz if and only if it is generated by two elements having orders 2 and 3, and whose product has order 7. This characterization has made possible much of the work in classifying Hurwitz groups. The most relevant for our investigation is the following result of Macbeath (see [M]):

> The simple group $PSL(2, q)$ is Hurwitz in exactly three cases:
> i) $q = 7$;
> ii) q is prime, with $q \equiv \pm 1 \pmod 7$;
> iii) $q = p^3$, with p prime and $p \equiv \pm 2, \pm 3 \pmod 7$.

In particular, $PSL(2, 8)$ and $PSL(2, 27)$ are Hurwitz groups. We shall require that q be relatively prime to $2 \cdot 3 \cdot 7$, but this excludes just three possibilities, namely $q \in \{7, 8, 27\}$. Note that in all of the cases we consider, $q \equiv \pm 1 \pmod 7$.

The order of $PSL(2, q)$ (for odd q) is $q(q^2 - 1)/2$. Hence we obtain

$$g = 1 + \frac{q(q^2 - 1)}{168}$$

as the genus of the corresponding curve(s).

For completeness, we remark that there are three distinct Hurwitz curves when q is prime (apart from $q = 7$), and just one when $q = p^3$. However, this has no bearing on the representations that we study.

In addition, there are other known families of Hurwitz groups. For example, all Ree groups are Hurwitz, as are all but finitely many of the alternating groups. See Conder [C] for a summary of such results.

2.2. **Ramification data.** Let X be a Hurwitz curve with automorphism group G and let

$$\psi : X \to Y = X/G \tag{3}$$

denote the quotient map. By again viewing X as a hyperbolic surface, the ramification data are easily deduced. The quotient Y is formed by one fundamental triangle and its mirror image, with the natural identifications on their boundaries. Hence it is a surface of genus 0 with 3 metric singularities. Thus ψ has exactly three branch points. The stabilizer subgroups of the corresponding ramification points in X are cyclic, of orders 2, 3, and 7. We label the three branch points P_1, P_2, and P_3, so that if $P \in \psi^{-1}(P_1)$, then P has stabilizer subgroup of order 2, if $P \in \psi^{-1}(P_2)$, P

has stabilizer subgroup of order 3, and if $P \in \psi^{-1}(P_3)$, P has stabilizer subgroup of order 7.

3. Representation theory of $PSL(2,q)$

3.1. **General theory on representations of PSL(2,q).** We first review the representation theory of $G = PSL(2,q)$ over $\mathbb{C}$, following the treatment in [FH], to fix notation.

Let $\mathbb{F} = GF(q)$ be the field with q elements. The group $PSL(2,q)$ has $3+(q-1)/2$ conjugacy classes of elements. Let $\varepsilon \in \mathbb{F}$ be a generator for the cyclic group $\mathbb{F}^\times$. Then each conjugacy class will have a representative of exactly one of the following forms:

$$\left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right), \left(\begin{array}{cc} x & 0 \\ 0 & x^{-1} \end{array}\right), \left(\begin{array}{cc} 1 & 1 \\ 0 & 1 \end{array}\right), \left(\begin{array}{cc} 1 & \varepsilon \\ 0 & 1 \end{array}\right), \left(\begin{array}{cc} x & \varepsilon y \\ y & x \end{array}\right). \tag{4}$$

The irreducible representations of $PSL(2,q)$ include the trivial representation $\mathbf{1}$ and one irreducible V of dimension q. All but two of the others fall into two types: representations W_α of dimension $q+1$ ("principal series"), and X_β of dimension $q-1$ ("discrete series"). The principal series representations W_α are indexed by homomorphisms $\alpha : \mathbb{F}^\times \to \mathbb{C}^\times$ with $\alpha(-1) = 1$. The discrete series representations X_β are indexed by homomorphisms $\beta : T \to \mathbb{C}^\times$ with $\beta(-1) = 1$, where T is a cyclic subgroup of order $q+1$ of $\mathbb{F}(\sqrt{\varepsilon})^\times$. The characters of these are as follows:

	$\left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right)$	$\left(\begin{array}{cc} x & 0 \\ 0 & x^{-1} \end{array}\right)$	$\left(\begin{array}{cc} 1 & 1 \\ 0 & 1 \end{array}\right)$	$\left(\begin{array}{cc} 1 & \varepsilon \\ 0 & 1 \end{array}\right)$	$\left(\begin{array}{cc} x & \varepsilon y \\ y & x \end{array}\right)$
$\mathbf{1}$	1	1	1	1	1
X_β	$q-1$	0	-1	-1	$-\beta(x+\sqrt{\varepsilon}y) - \beta(x-\sqrt{\varepsilon}y)$
V	q	1	0	0	-1
W_α	$q+1$	$\alpha(x)+\alpha(x^{-1})$	1	1	0

Let ζ be a primitive qth root of unity in $\mathbb{C}$. Let ξ and ξ' be defined by

$$\xi = \sum_{(\frac{a}{q})=1} \zeta^a \text{ and } \xi' = \sum_{(\frac{a}{q})=-1} \zeta^a, \tag{5}$$

where the sums are over the quadratic residues and nonresidues $\pmod q$, respectively. If $q \equiv 1 \bmod 4$, then the principal series representation W_{α_0} corresponding to

$$\begin{array}{rcl} \alpha_0 : \mathbb{F}^\times & \to & \mathbb{C}^\times \\ \varepsilon & \mapsto & -1 \end{array}$$

is not irreducible, but splits into two irreducibles W' and W'', each of dimension $(q+1)/2$. Their characters satisfy:

	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} x & 0 \\ 0 & x^{-1} \end{pmatrix}$	$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & \varepsilon \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} x & \varepsilon y \\ y & x \end{pmatrix}$
W'	$\frac{q+1}{2}$	$\alpha_0(x)$	$1+\xi$	$1+\xi'$	0
W''	$\frac{q+1}{2}$	$\alpha_0(x)$	$1+\xi'$	$1+\xi$	0

Let τ denote a generator of T. Similarly, if $q \equiv 3 \bmod 4$, then the discrete series representation X_{β_0} corresponding to

$$\begin{array}{rcl} \beta_0 : T & \to & \mathbb{C}^\times \\ \tau & \mapsto & -1 \end{array}$$

splits into two irreducibles X' and X'', each of dimension $(q-1)/2$. Their characters satisfy:

	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} x & 0 \\ 0 & x^{-1} \end{pmatrix}$	$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & \varepsilon \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} x & \varepsilon y \\ y & x \end{pmatrix}$
X'	$\frac{q-1}{2}$	0	ξ	ξ'	$-\beta_0(x+y\sqrt{\varepsilon})$
X''	$\frac{q-1}{2}$	0	ξ'	ξ	$-\beta_0(x+y\sqrt{\varepsilon})$

According to Janusz [Ja], the Schur index of each irreducible representation of G is 1.

There is a "Galois action" on the set of equivalence classes of irreducible representations of G as follows. Let χ denote an irreducible character. The character values $\chi(g)$ lie in $\mathbb{Q}(\mu)$, where μ is a primitive m^{th} root of unity and $m = q(q^2-1)/4$. Let $\mathcal{G} = Gal(\mathbb{Q}(\mu)/\mathbb{Q})$ denote the Galois group. For each integer j relatively prime to m, there is an element σ_j of $\mathcal{G}$ taking μ to μ^j. This Galois group element will act on representations by taking a representation with character values $(a_1, \ldots, a_n)$ to a representation with character values $(\sigma_j(a_1), \ldots, \sigma_j(a_n))$. Representations with rational character values will be fixed under this action. Because the Schur index of each representation is 1, representations with rational character values will be defined over $\mathbb{Q}$.

The action of the Galois group $\mathcal{G}$ can easily be seen from the character table. It will fix the trivial representation and the q-dimensional representation V. Its action permutes the set of $q-1$-dimensional "principal series" representations X_β, and the set of $q+1$-dimensional "discrete series" representations W_α. In the case $q \equiv 1 \pmod 4$, the Galois group will exchange the two $(q+1)/2$-dimensional representations W' and W''; if $q \equiv 3 \pmod 4$, the Galois group will exchange the two $(q-1)/2$-dimensional representations X' and X''.

3.2. **Induced characters.** We will be interested in the induced characters from subgroups of orders 2, 3, and 7. For each value of q, each of these subgroups is unique up to conjugacy; we can choose subgroups H_2 of order 2, H_3 of order 3, and H_7 of order 7 that are generated by elements of the form

$$\begin{pmatrix} x & 0 \\ 0 & x^{-1} \end{pmatrix} \text{ or } \begin{pmatrix} x & \varepsilon y \\ y & x \end{pmatrix}.$$

Which of these two forms each generator will take depends on q mod 4, mod 3, and mod 7, respectively. Recall that we defined generators ε of the cyclic group $\mathbb{F}^\times$, of order $q-1$, and τ of the cyclic group $T \subseteq \mathbb{F}(\sqrt{\varepsilon})^\times$ of order $q+1$, respectively. We define numbers i, ω, and ϕ to be primitive roots of unity as follows.

When $q \equiv 1 \pmod 4$, let i denote an element in $\mathbb{F}^\times$ whose square is -1 (one can take $i = \varepsilon^{(q-1)/4}$). Then the subgroup H_2 of order 2 in $PSL(2,q)$ is generated by

$$\begin{pmatrix} i & 0 \\ 0 & i^{-1} \end{pmatrix}.$$

If $q \equiv 3 \pmod 4$, then we take $i = x_i + \sqrt{\varepsilon} y_i$ to be an element of T whose square is -1 (one can take $i = \tau^{(q+1)/4}$). Then the subgroup H_2 of order 2 in $PSL(2,q)$ is generated by

$$\begin{pmatrix} x_i & \varepsilon y_i \\ y_i & x_i \end{pmatrix}.$$

Similarly, we define ω to be a primitive 6th root of unity. In the case where $q \equiv 1 \pmod 3$, we can take $\omega = \varepsilon^{(q-1)/6} \in \mathbb{F}^\times$. When $q \equiv -1 \pmod 3$, we take $\omega = x_\omega + \sqrt{\varepsilon} y_\omega = \tau^{(q+1)/6} \in T$. The subgroup H_3 of order 3 in $PSL(2,q)$ will then be generated by

$$\begin{pmatrix} \omega & 0 \\ 0 & \omega^{-1} \end{pmatrix}, \text{ if } q \equiv 1 \pmod 3, \text{ or } \begin{pmatrix} x_\omega & \varepsilon y_\omega \\ y_\omega & x_\omega \end{pmatrix}, \text{ if } q \equiv -1 \pmod 3.$$

Lastly, we want to define ϕ to be a primitive 14th root of unity. Recall that $q \equiv \pm 1 \pmod 7$. If $q \equiv 1 \pmod 7$, then we can take $\phi = \varepsilon^{(q-1)/14} \in \mathbb{F}^\times$, and if $q \equiv -1 \pmod 7$, then we can take $\phi = x_\phi + \sqrt{\varepsilon} y_\phi = \tau^{(q+1)/14} \in T$. The subgroup H_7 of order 7 in $PSL(2,q)$ will then be generated by

$$\begin{pmatrix} \phi & 0 \\ 0 & \phi^{-1} \end{pmatrix}, \ q \equiv 1 \pmod 3, \text{ or } \begin{pmatrix} x_\phi & \varepsilon y_\phi \\ y_\phi & x_\phi \end{pmatrix}, \ q \equiv -1 \pmod 3.$$

With these definitions, it is easy to compute the restrictions of the irreducible representations of $PSL(2,q)$ to the subgroups above. We omit the details, but the computations for the groups of order 2 and 3 are given in [JK2], and the computation for the group of order 7 is very similar. Using Frobenius reciprocity, we then obtain the corresponding induced representations. In each case, we denote a primitive character of the cyclic group H_k by θ_k.

3.2.1. *Induced characters from H_2.* The induced representations from the nontrivial character of H_2 are given below. The multiplicities depend on q (mod 8). Note that most representation have the same multiplicity as V. When $i \in \mathbb{F}^\times$, i.e. when $q \equiv 1 \pmod 4$, the multiplicity of a discrete series representation W_α depends on the sign of $\alpha(i)$. Recall that $\alpha(-1) = 1$, so $\alpha(i) = \pm 1$. the multiplicity of W_α will be the same as the multiplicity of V if $\alpha(i) = 1$ and one larger if $\alpha(i) = -1$. Similarly, when $q \equiv 3 \pmod 4$ and $i \in T$, the multiplicity of a principal series representation X_β depends on the sign of $\beta(i)$. In this case the multiplicity of X_β will be the same as the multiplicity of V when $\beta(i) = 1$, and one less if $\beta(i) = -1$. Lastly, the signs of $\alpha_0(i)$ or $\beta_0(i)$ depend on $q \pmod 8$ and determine the multiplicities of W' and W'' or X' and X'', respectively. A similar pattern will hold for the induced representations from H_3 and H_7.

For $q \equiv 1 \pmod 8$,

$$Ind_{H_2}^G \theta_2 = \frac{q-1}{2}\left[\frac{1}{2}(W'+W'') + \sum_\beta X_\beta + V + \sum_{\alpha(i)=1} W_\alpha\right] + \frac{q+3}{2}\sum_{\alpha(i)=-1} W_\alpha.$$

For $q \equiv 3 \pmod 8$,

$$Ind_{H_2}^G \theta_2 = \frac{q+1}{2}\left[\sum_{\beta(i)=1} X_\beta + V + \sum_\alpha W_\alpha\right] + \frac{q-3}{2}\left[\frac{1}{2}(X'+X'') + \sum_{\beta(i)=-1} X_\beta\right].$$

For $q \equiv 5 \pmod 8$,

$$Ind_{H_2}^G \theta_2 = \frac{q-1}{2}\left[\sum_\beta X_\beta + V + \sum_{\alpha(i)=1} W_\alpha\right] + \frac{q+3}{2}\left[\frac{1}{2}(W'+W'') + \sum_{\alpha(i)=-1} W_\alpha\right].$$

And for $q \equiv 7 \pmod 8$,

$$Ind_{H_2}^G \theta_2 = \frac{q+1}{2}\left[\frac{1}{2}(X'+X'') + \sum_{\beta(i)=1} X_\beta + V + \sum_\alpha W_\alpha\right] + \frac{q-3}{2}\sum_{\beta(i)=-1} X_\beta.$$

3.2.2. *Induced characters from H_3.* The induced representations from the two nontrivial characters θ_3 and θ_3^2 of H_3 are the same. In this case the multiplicities depend on $q \pmod{12}$, which determines whether the 6th root of unity ω is in $\mathbb{F}^\times$, or in $T \subset \mathbb{F}(\sqrt{\varepsilon})^\times$. Now the multiplicity of a discrete (resp. principal) series representation W_α (resp. X_β) will be the same as the multiplicity of V if $\alpha(\phi) = 1$ (resp. $\beta(\phi) = 1$) and one larger (resp. smaller) if $\alpha(\phi) = e^{\frac{\pm 2\pi i}{3}}$ (resp. $\beta(\phi) = e^{\frac{\pm 2\pi i}{3}}$). The signs of $\alpha_0(\omega)$ or $\beta_0(\omega)$ depend on $q \pmod{12}$ and determine the multiplicities of W' and W'' or X' and X'', respectively.

If $q \equiv 1 \pmod{12}$, we have

$$Ind_{H_3}^G \theta_3 = \frac{q-1}{3}\left[\frac{1}{2}(W'+W'') + \sum_\beta X_\beta + V + \sum_{\alpha(\omega)=1} W_\alpha\right] + \frac{q+2}{3}\sum_{\alpha(\omega)=e^{\frac{\pm 2\pi i}{3}}} W_\alpha.$$

If $q \equiv 5 \pmod{12}$, we have

$$Ind_{H_3}^G \theta_3 = \frac{q+1}{3}\left[\frac{1}{2}(W'+W'') + \sum_{\beta(\omega)=1} X_\beta + V + \sum_\alpha W_\alpha\right] + \frac{q-2}{3}\sum_{\beta(\omega)=1} X_\beta.$$

If $q \equiv 7 \pmod{12}$, we have

$$Ind_{H_3}^G \theta_3 = \frac{q-1}{3}\left[\frac{1}{2}(X'+X'') + \sum_\beta X_\beta + V + \sum_{\alpha(\omega)=1} W_\alpha\right] + \frac{q+2}{3}\sum_{\alpha(\omega)=e^{\frac{\pm 2\pi i}{3}}} W_\alpha.$$

And if $q \equiv 11 \pmod{12}$, we have

$$Ind_{H_3}^G \theta_3 = \frac{q+1}{3}\left[\frac{1}{2}(X'+X'') + \sum_{\beta(\omega)=1} X_\beta + V + \sum_\alpha W_\alpha\right] + \frac{q-2}{3}\sum_{\beta(\omega)=e^{\frac{\pm 2\pi i}{3}}} X_\beta.$$

3.2.3. *Induced characters from H_7.* For H_7, the induced representations from the six nontrivial characters θ_7^k are not all the same, but depend on k. These representations also depend on $q \pmod{28}$, which determines whether the 14th root of unity ϕ is in $\mathbb{F}^\times$ or $\mathbb{F}(\sqrt{\varepsilon})^\times$. For an induced nontrivial character $\mathrm{Ind}_{H_7}^G \theta_7^k$, the multiplicity of a discrete (resp. principal) series representation W_α (resp. X_β) will be the same as the multiplicity of V if $\alpha(\phi) \neq e^{\pm \frac{2\pi i k}{7}}$ (resp. $\beta(\phi) \neq e^{\pm \frac{2\pi i k}{7}}$) and one larger (resp. smaller) if $\alpha(\phi) = e^{\pm \frac{2\pi i k}{7}}$ (resp. $\beta(\phi) = e^{\pm \frac{2\pi i k}{7}}$). The signs of $\alpha_0(\phi)$ or $\beta_0(\phi)$ depend on $q \pmod{28}$ and determine the multiplicities of W' and W'' or X' and X'', respectively.

If $q \equiv 1 \pmod{28}$, we have

$$Ind_{H_7}^G \theta_7^k = \frac{q-1}{7}\left[\frac{1}{2}(W'+W'') + \sum_\beta X_\beta + V + \sum_{\alpha(\phi)\neq e^{\pm\frac{2\pi i k}{7}}} W_\alpha\right] + \frac{q+6}{7}\sum_{\alpha(\phi)=e^{\pm\frac{2\pi i k}{7}}} W_\alpha.$$

If $q \equiv 13 \pmod{28}$, we have

$$Ind_{H_7}^G \theta_7^k = \frac{q+1}{7}\left[\frac{1}{2}(W'+W'') + \sum_{\beta(\phi)\neq e^{\pm\frac{2\pi i k}{7}}} X_\beta + V + \sum_\alpha W_\alpha\right] + \frac{q-6}{7}\sum_{\beta(\phi)=e^{\pm\frac{2\pi i k}{7}}} X_\beta.$$

If $q \equiv 15 \pmod{28}$, we have

$$Ind_{H_7}^G \theta_7^k = \frac{q-1}{7}\left[\frac{1}{2}(X'+X'') + \sum_\beta X_\beta + V + \sum_{\alpha(\phi)\neq e^{\pm\frac{2\pi i k}{7}}} W_\alpha\right] + \frac{q+6}{7}\sum_{\alpha(\phi)=e^{\pm\frac{2\pi i k}{7}}} W_\alpha.$$

And if $q \equiv 27 \pmod{28}$, we have

$$Ind_{H_7}^G \theta_7^k = \frac{q+1}{7}\left[\frac{1}{2}(X'+X'') + \sum_{\beta(\phi)\neq e^{\pm\frac{2\pi i k}{7}}} X_\beta + V + \sum_\alpha W_\alpha\right] + \frac{q-6}{7}\sum_{\beta(\phi)=e^{\pm\frac{2\pi i k}{7}}} X_\beta.$$

4. The Riemann-Roch space as a G-module

Now we have all of the pieces we need to compute the G-module structure of the Riemann-Roch space $L(D)$ of a general G-invariant divisor D. We will first compute the ramification module, which does not depend on D. We will then compute the equivariant degree of D, and use the equivariant Riemann-Roch formula (1) to compute $L(D)$.

4.1. **Ramification module.** The ramification module introduced by Kani [K] and Nakajima [N] is defined by

$$\Gamma_G = \sum_{P \in X_{\mathrm{ram}}} \mathrm{Ind}_{G_P}^G \left(\sum_{\ell=1}^{e_P - 1} \ell \theta_P^\ell \right),$$

where the first sum is over the ramification points of $\psi : X \to Y = X/G$, and θ_P is the ramification character at a point P. Both Kani and Nakajima showed that

there is a G-module $\tilde{\Gamma}_G$ such that $\Gamma_G \simeq \bigoplus_{|G|} \tilde{\Gamma}_G$. Because Γ_G does not figure in our calculations, we abuse notation and refer to $\tilde{\Gamma}_G$ as the *ramification module*.

Recall from section 2.2 that $\psi : X \to Y = X/G$ has three branch points, P_1, P_2, and P_3. If $P \in \psi^{-1}(P_1)$, G_P has order 2, so there are $\frac{|G|}{2}$ ramification points where G_P is conjugate to H_2. If $P \in \psi^{-1}(P_2)$, G_P has order 3, so there are $\frac{|G|}{3}$ ramification points where G_P is conjugate to H_3, and if $P \in \psi^{-1}(P_3)$, G_P has order 7, so there are $\frac{|G|}{7}$ ramification points where G_P is conjugate to H_7. Thus

$$\tilde{\Gamma}_G = \frac{1}{|G|}\left(\frac{|G|}{2}\operatorname{Ind}_{H_2}^G \theta_2 + \frac{|G|}{3}\sum_{\ell=1}^{2} \ell \operatorname{Ind}_{H_3}^G \theta_3^\ell + \frac{|G|}{7}\sum_{\ell=1}^{6} \ell \operatorname{Ind}_{H_7}^G \theta_7^\ell\right). \tag{6}$$

To compute this, we break it into three pieces:

$$\begin{aligned}
\tilde{\Gamma}_G &= \Gamma_{H_2} + \Gamma_{H_3} + \Gamma_{H_7},\\
\Gamma_{H_2} &= \frac{1}{2}\operatorname{Ind}_{H_2}^G \theta_2,\\
\Gamma_{H_3} &= \frac{1}{3}(\operatorname{Ind}_{H_3}^G \theta_3 + 2\operatorname{Ind}_{H_3}^G \theta_3^2),\\
\Gamma_{H_7} &= \frac{1}{7}(\operatorname{Ind}_{H_7}^G \theta_7 + 2\operatorname{Ind}_{H_7}^G \theta_7^2 + 3\operatorname{Ind}_{H_7}^G \theta_7^3\\
&\quad +4\operatorname{Ind}_{H_7}^G \theta_7^4 + 5\operatorname{Ind}_{H_7}^G \theta_7^5 + 6\operatorname{Ind}_{H_7}^G \theta_7^6).
\end{aligned}$$

Each piece is then computed from the induced characters in section 3.2. Γ_{H_2} depends on $q \pmod 8$.

For $q \equiv 1 \pmod 8$,

$$\Gamma_{H_2} = \frac{q-1}{4}\left[\frac{1}{2}(W' + W'') + \sum_{\beta} X_\beta + V + \sum_{\alpha(i)=1} W_\alpha\right] + \frac{q+3}{4}\sum_{\alpha(i)=-1} W_\alpha.$$

For $q \equiv 3 \pmod 8$,

$$\Gamma_{H_2} = \frac{q+1}{4}\left[\sum_{\beta(i)=1} X_\beta + V + \sum_{\alpha} W_\alpha\right] + \frac{q-3}{4}\left[\frac{1}{2}(X' + X'') + \sum_{\beta(i)=-1} X_\beta\right].$$

For $q \equiv 5 \pmod 8$,

$$\Gamma_{H_2} = \frac{q-1}{4}\left[\sum_{\beta} X_\beta + V + \sum_{\alpha(i)=1} W_\alpha\right] + \frac{q+3}{4}\left[\frac{1}{2}(W' + W'') + \sum_{\alpha(i)=-1} W_\alpha\right].$$

And for $q \equiv 7 \pmod 8$,

$$\Gamma_{H_2} = \frac{q+1}{4}\left[\frac{1}{2}(X' + X'') + \sum_{\beta(i)=1} X_\beta + V + \sum_{\alpha} W_\alpha\right] + \frac{q-3}{4}\sum_{\beta(i)=-1} X_\beta.$$

The contribution Γ_{H_3} of H_3 to the ramification module is

$$\Gamma_{H_3} = \frac{1}{3}\left(\operatorname{Ind}_{H_3}^G \theta_3 + 2\operatorname{Ind}_{H_3}^G \theta_3^2\right) = \operatorname{Ind}_{H_3}^G \theta_3,$$

since $\operatorname{Ind}_{H_3}^G \theta_3$ and $\operatorname{Ind}_{H_3}^G \theta_3^2$ are the same. This character was computed in section 3.2.

For H_7, the induced representations from the six nontrivial characters θ_7^k are not all the same. However, the representations $\mathrm{Ind}_{H_7}^G \theta_7^k$ and $\mathrm{Ind}_{H_7}^G \theta_7^{-k}$ are equal. Thus Γ_{H_7} is

$$\begin{aligned}\Gamma_{H_7} &= \frac{1}{7}\left(\mathrm{Ind}_{H_7}^G \theta_7 + 2\,\mathrm{Ind}_{H_7}^G \theta_7^2 + \ldots + 6\,\mathrm{Ind}_{H_7}^G \theta_7^6\right) \\ &= \frac{1}{7}\left(7\,\mathrm{Ind}_{H_7}^G \theta_7 + 7\,\mathrm{Ind}_{H_7}^G \theta_7^2 + 7\,\mathrm{Ind}_{H_7}^G \theta_7^4\right) \\ &= \mathrm{Ind}_{H_7}^G \theta_7 + \mathrm{Ind}_{H_7}^G \theta_7^2 + \mathrm{Ind}_{H_7}^G \theta_7^4.\end{aligned}$$

Recall from section 3.2 that the multiplicities of the irreducible representations W_α and X_β in the induced representation $\mathrm{Ind}_{H_7}^G \theta_7^k$ depend on the value of $\alpha(\phi)$ or $\beta(\phi)$, and that this value must be $e^{\frac{2\pi i k}{7}}$ for some $k = 0, \ldots, 6$. In the sum $\Gamma_{H_7} = \mathrm{Ind}_{H_7}^G \theta_7 + \mathrm{Ind}_{H_7}^G \theta_7^2 + \mathrm{Ind}_{H_7}^G \theta_7^4$ we will have, for example for the multiplicities of the W_α when $q \equiv 1 \pmod{28}$,

$$\begin{aligned}\Gamma_{H_7} &= \mathrm{Ind}_{H_7}^G \theta_7 + \mathrm{Ind}_{H_7}^G \theta_7^2 + \mathrm{Ind}_{H_7}^G \theta_7^4 \\ &= \frac{q-1}{7} \sum_{\alpha(\phi) \neq e^{\pm \frac{2\pi i}{7}}} W_\alpha + \frac{q+6}{7} \sum_{\alpha(\phi) = e^{\pm \frac{2\pi i}{7}}} W_\alpha \\ &+ \frac{q-1}{7} \sum_{\alpha(\phi) \neq e^{\pm \frac{4\pi i}{7}}} W_\alpha + \frac{q+6}{7} \sum_{\alpha(\phi) = e^{\pm \frac{4\pi i}{7}}} W_\alpha \\ &+ \frac{q-1}{7} \sum_{\alpha(\phi) \neq e^{\pm \frac{8\pi i}{7}}} W_\alpha + \frac{q+6}{7} \sum_{\alpha(\phi) = e^{\pm \frac{8\pi i}{7}}} W_\alpha \\ &+ \text{ other characters.}\end{aligned}$$

This adds up to

$$\Gamma_{H_7} = \frac{3q+4}{7} \sum_{\alpha(\phi) \neq 1} W_\alpha + \frac{3q-3}{7} \sum_{\alpha(\phi)=1} W_\alpha + \text{ other characters.}$$

The multiplicities of the other irreducible characters in $\mathrm{Ind}_{H_7}^G \theta_7^k$ do not depend on k. Adding these in, the total for the case $q \equiv 1 \pmod{28}$ is

$$\Gamma_{H_7} = \frac{3q-3}{7}\left[\sum_\beta X_\beta + V + \sum_{\alpha(\phi)=1} W_\alpha + \frac{1}{2}(W' + W'')\right] + \frac{3q+4}{7} \sum_{\alpha(\phi)\neq 1} W_\alpha.$$

Similar calculations yield the following. If $q \equiv 13 \pmod{28}$,

$$\Gamma_{H_7} = \frac{3q+3}{7}\left[\sum_{\beta(\phi)=1} X_\beta + V + \sum_\alpha W_\alpha + \frac{1}{2}(W' + W'')\right] + \frac{3q-4}{7} \sum_{\beta(\phi)\neq 1} X_\beta.$$

If $q \equiv 15 \pmod{28}$, we have

$$\Gamma_{H_7} = \frac{3q-3}{7}\left[\sum_{\beta} X_\beta + V + \sum_{\alpha(\phi)=1} W_\alpha + \frac{1}{2}(X' + X'')\right] + \frac{3q+4}{7}\sum_{\alpha(\phi)\neq 1} W_\alpha.$$

And if $q \equiv 27 \pmod{28}$, we have

$$\Gamma_{H_7} = \frac{3q+3}{7}\left[\sum_{\beta(\phi)=1} X_\beta + V + \sum_{\alpha} W_\alpha + \frac{1}{2}(X' + X'')\right] + \frac{3q-4}{7}\sum_{\beta(\phi)\neq 1} X_\beta.$$

To compute the ramification module, we sum the components Γ_{H_2}, Γ_{H_3}, and Γ_{H_7} listed above. The following numbers will be useful.

Definition 1. *For each possible equivalence class of $q \pmod{84}$, we define a* ***base multiplicity*** *m, as follows:*

- *If $q \equiv 1, 13, 29,$ or $43 \pmod{84}$, then $m = q + \lfloor \frac{q}{84} \rfloor$.*
- *If $q \equiv 41, 55, 71,$ or $83 \pmod{84}$, then $m = q + \lceil \frac{q}{84} \rceil$.*

Definition 2. *Let $\alpha : \mathbb{F}^\times \to \mathbb{C}^\times$ be a character of $\mathbb{F}^\times$. Then we define a number*

$$N_\alpha = \#\{x \in \{i, \omega, \phi\} \mid x \in \mathbb{F}^\times \text{ and } \alpha(x) \neq 1\}.$$

Definition 3. *Recall that T is the cyclic subgroup of $\mathbb{F}(\sqrt{\varepsilon})^\times$ of order $q+1$. Let $\beta : T \to \mathbb{C}^\times$ be a character of T. Then we define a number*

$$N_\beta = \#\{x \in \{i, \omega, \phi\} \mid x \in T \text{ and } \beta(x) \neq 1\}.$$

Theorem 4. *We have the following decomposition of the ramification module:*

- *If $q \equiv 1 \pmod{8}$, then*

$$\tilde{\Gamma}_G = \frac{m}{2}(W' + W'') + mV + \sum_{\beta}(m - N_\beta)X_\beta + \sum_{\alpha}(m + N_\alpha)W_\alpha$$

- *If $q \equiv 3 \pmod{8}$, then*

$$\tilde{\Gamma}_G = \frac{m-1}{2}(X' + X'') + mV + \sum_{\beta}(m - N_\beta)X_\beta + \sum_{\alpha}(m + N_\alpha)W_\alpha$$

- *If $q \equiv 5 \pmod{8}$, then*

$$\tilde{\Gamma}_G = \frac{m+1}{2}(W' + W'') + mV + \sum_{\beta}(m - N_\beta)X_\beta + \sum_{\alpha}(m + N_\alpha)W_\alpha$$

- *If $q \equiv 7 \pmod{8}$, then*

$$\tilde{\Gamma}_G = \frac{m}{2}(X' + X'') + mV + \sum_{\beta}(m - N_\beta)X_\beta + \sum_{\alpha}(m + N_\alpha)W_\alpha$$

4.2. **Equivariant degree.** Now we will define and compute the equivariant degree of a G-invariant divisor. (See for example [B] for more details). This, together with the equivariant Riemann-Roch formula (1), will allow us to compute the G-module structure of the Riemann-Roch space $L(D)$.

Fix a point $P \in X$ and let D be a divisor on X of the form

$$D = \frac{1}{e_P} \sum_{g \in G} g(P) = \sum_{g \in G/G_P} g(P),$$

where G_P denotes the stabilizer in G of P and $e_P = |G_P|$ denotes the ramification index at P. Such a divisor is called a *reduced orbit*; any G-invariant divisor on X can be written as a sum of multiples of reduced orbits.

The *equivariant degree* of a multiple rD of a reduced orbit is the virtual representation

$$\deg_{eq}(rD) = \begin{cases} \operatorname{Ind}_{G_P}^G \sum_{\ell=1}^{r} \theta_P^{-\ell}, & r > 0 \\ 0, & r = 0 \\ -\operatorname{Ind}_{G_P}^G \sum_{\ell=0}^{|r|-1} \theta_P^{-\ell}, & r < 0 \end{cases}$$

where θ_P is the ramification character of X at P (a nontrivial character of G_P). In general, the equivariant degree is additive on disjointly supported divisors. Note that if r is a multiple of e_P, then then D is the pull-back of a divisor on X/G via ψ in (3), and the equivariant degree is a multiple of the regular representation $\mathbb{C}[G]$ of G. More generally, if D is a reduced orbit and $r = e_P r' + r''$, then

$$\deg_{eq}(rD) = r' \cdot \mathbb{C}[G] + \deg_{eq}(r''D).$$

(Note this is true even when r' is negative).

On the Hurwitz curve X, the results of section 2.2 tell us that there are only four types of reduced orbits to consider: the stabilizer G_P of a point P in the support of D may have order 1, 2, 3, or 7, and therefore be either trivial or conjugate to H_2, H_3, or H_7. Let D_1, D_2, D_3, and D_7 denote reduced orbits of each type. There is only one choice of reduced orbit for D_2, D_3, and D_7; for D_1 we see from the definition that the equivariant degree does not depend on our choice of orbit. Given a point in D_1, the stabilizer is trivial, so the divisor is a pullback and the equivariant degree is

$$\deg_{eq}(D_1) = \mathbb{C}[G].$$

A general G-invariant divisor may be written as $r_1 D_1 + r_2 D_2 + r_3 D_3 + r_7 D_7$. If we write $r_2 = 2r_2' + r_2''$, $r_3 = 3r_3' + r_3''$, and $r_7 = 7r_7' + r_7''$, then we have

$$\begin{aligned} &\deg_{eq}(r_1 D_1 + r_2 D_2 + r_3 D_3 + r_7 D_7) \\ &\quad = \deg_{eq}((r_1 + r_2' + r_3' + r_7')D_1 + r_2'' D_2 + r_3'' D_3 + r_7'' D_7) \\ &\quad = (r_1 + r_2' + r_3' + r_7')\mathbb{C}[G] + \deg_{eq}(r_2'' D_2 + r_3'' D_3 + r_7'' D_7). \end{aligned}$$

Therefore, to compute the equivariant degree of a general divisor, all that remains is to compute $\deg_{eq}(r_i D_i)$ for $i \in \{2, 3, 7\}$, where we may assume that $1 \leq r_i < i$.

Case 1: : $r_2 D_2$. Given our assumptions, the only possibility is that $r_2 = 1$. Given a point P in the support of D_2, the stabilizer G_P is conjugate to H_2. In this case, the equivariant degree of D_2 is

$$\deg_{eq}(D_2) = \mathrm{Ind}_{H_2}^G \theta_2.$$

Case 2: : $r_3 D_3$. Here we may have either $r_3 = 1$ or $r_3 = 2$. The stabilizer of a point in the support of D_3 is conjugate to H_3. Recall that $\mathrm{Ind}_{H_3}^G \theta_3^2 = \mathrm{Ind}_{H_3}^G \theta_3$, so we have

$$\begin{aligned} \deg_{eq}(D_2) &= \mathrm{Ind}_{H_3}^G \theta_3 \\ \deg_{eq}(2D_2) &= 2\,\mathrm{Ind}_{H_3}^G \theta_3. \end{aligned}$$

Case 3: : $r_7 D_7$. In this case, we have $1 \le r_7 \le 6$. The stabilizer of a point in the support of D_7 is conjugate to H_7. Recall that for $k = 1, \ldots, 6$, $\mathrm{Ind}_{H_7}^G \theta_7^k = \mathrm{Ind}_{H_7}^G \theta_7^{-k}$. Therefore the equivariant degree is as follows:

- $\deg_{eq}(D_7) = \mathrm{Ind}_{H_7}^G \theta_7$.
- $\deg_{eq}(2D_7) = \mathrm{Ind}_{H_7}^G \theta_7 + \mathrm{Ind}_{H_7}^G \theta_7^2$.
- $\deg_{eq}(3D_7) = \mathrm{Ind}_{H_7}^G \theta_7 + \mathrm{Ind}_{H_7}^G \theta_7^2 + \mathrm{Ind}_{H_7}^G \theta_7^3$, which is the same as the H_7 component of the ramification module, Γ_{H_7}.
- $\deg_{eq}(4D_7) = \Gamma_{H_7} + \mathrm{Ind}_{H_7}^G \theta_7^3$.
- $\deg_{eq}(5D_7) = \Gamma_{H_7} + \mathrm{Ind}_{H_7}^G \theta_7^3 + \mathrm{Ind}_{H_7}^G \theta_7^2$.
- $\deg_{eq}(6D_7) = 2\Gamma_{H_7}$.

Now we add these up. As in the case of the ramification module, the equivariant degree is most conveniently written in terms of a "base multiplicity" and modifiers. We define the base multiplicity as follows.

- If $q \equiv 1 \pmod 4$, then let $b_2 = r_2\left(\frac{q-1}{2}\right)$. Otherwise, if $q \equiv 3 \pmod 4$, then let $b_2 = r_2\left(\frac{q+1}{2}\right)$.
- If $q \equiv 1 \pmod 3$, then let $b_3 = r_3\left(\frac{q-1}{3}\right)$, and if $q \equiv 2 \pmod 3$, then let $b_3 = r_3\left(\frac{q+1}{3}\right)$.
- Similarly, if $q \equiv 1 \pmod 7$, then let $b_7 = r_7\left(\frac{q-1}{7}\right)$, and if $q \equiv 6 \pmod 7$, then let $b_7 = r_7\left(\frac{q+1}{7}\right)$.

The base multiplicity is then defined to be

$$\begin{aligned} b &= b_2 + b_3 + b_7 \\ &= r_2\left(\frac{q \pm 1}{2}\right) + r_3\left(\frac{q \pm 1}{3}\right) + r_7\left(\frac{q \pm 1}{7}\right). \end{aligned}$$

Then the equivariant degree $\deg_{eq}(D)$ of the divisor $D = r_1 D_1 + r_2 D_2 + r_3 D_3 + r_7 D_7$, with $0 \le r_2 \le 1$, $0 \le r_3 \le 2$, and $0 \le r_7 \le 6$, is

$$\deg_{eq}(D) = b\left[\sum_{\beta} X_\beta + V + \sum_{\alpha} W_\alpha\right] + \text{ modifiers}, \tag{7}$$

where the modifiers are listed in the table below. For each q, three of the rows below will be added.

q	Modifiers to equivariant degree
$q \equiv 1 \pmod 8$	$+ r_2 \sum_{\alpha(i)=-1} W_\alpha + \frac{b}{2}(W' + W'')$
$q \equiv 3 \pmod 8$	$- r_2 \sum_{\beta(i)=-1} X_\beta + \frac{b - r_2}{2}(X' + X'')$
$q \equiv 5 \pmod 8$	$+ r_2 \sum_{\alpha(i)=-1} W_\alpha + \frac{b + r_2}{2}(W' + W'')$
$q \equiv 7 \pmod 8$	$- r_2 \sum_{\beta(i)=-1} X_\beta + \frac{b}{2}(X' + X'')$
$q \equiv 1 \pmod 3$	$+ r_3 \sum_{\alpha(\omega)\neq 1} W_\alpha$
$q \equiv 2 \pmod 3$	$-r_3 \sum_{\beta(\omega)\neq 1} X_\beta$
$q \equiv 1 \pmod 7$	$+ \sum_{k=1}^{r_7} \sum_{\alpha(\phi)=e^{\pm\frac{2\pi i k}{7}}} W_\alpha$
$q \equiv 6 \pmod 7$	$- \sum_{k=1}^{r_7} \sum_{\beta(\phi)=e^{\pm\frac{2\pi i k}{7}}} X_\beta$

4.3. **The Riemann-Roch space.** Now we would like to compute the G-module structure of the Riemann-Roch space $L(D)$ for a G-invariant divisor D. First, let us consider which G-invariant divisors are non-special. To be non-special, it is sufficient to have $\deg D > 2g - 2$, where

$$g = 1 + \frac{(q)(q^2 - 1)}{168}$$

is the genus of X, so $2g - 2 = \frac{1}{84}q(q^2 - 1) = \frac{1}{168}|G|$. The reduced orbits D_1, D_2, D_3 and D_7 have degrees $|G|$, $|G|/2$, $|G|/3$, and $|G|/7$, respectively. Therefore if a G-invariant divisor $r_1D_1 + r_2D_2 + r_3D_3 + r_7D_7$ has positive degree, the smallest its degree could be is $|G|/42$, which is strictly larger than $2g - 2$. Therefore any G-invariant divisor with positive degree is non-special.

Thus for any G-invariant divisor D with positive degree, we may use the equivariant Riemann-Roch formula (1) to compute the G-module structure of the Riemann-Roch space $L(D)$:

$$[L(D)] = (1 - g_{X/G})[\mathbb{C}[G]] + [\deg_{eq}(D)] - [\tilde{\Gamma}_G].$$

Since $X/G \cong \mathbb{P}^1$, its genus is zero. As in section 4.2, we may assume that $D = r_1D_1+r_2D_2+r_3D_3+r_7D_7$, with $0 \le r_2 \le 1$, $0 \le r_3 \le 2$, and $0 \le r_7 \le 6$. Combining the results and notation of sections 4.1 and 4.2, we obtain the following.

$$L(D) = (1+r_1)\mathbb{C}[G] + (b-m)\left[\sum_\beta X_\beta + V + \sum_\alpha W_\alpha\right] + \text{ modifiers,}$$

where the modifiers depend on $q \pmod{168}$ and are listed in the following table. Again, for each value of q, three of the rows below will be added.

q	Modifiers to Riemann-Roch space
$q \equiv 1 \pmod 8$	$+ (r_2-1)\sum_{\alpha(i)=-1} W_\alpha + \frac{b-m}{2}(W'+W'')$
$q \equiv 3 \pmod 8$	$+ (1-r_2)\sum_{\beta(i)=-1} X_\beta + \frac{b-m+1-r_2}{2}(X'+X'')$
$q \equiv 5 \pmod 8$	$+ (r_2-1)\sum_{\alpha(i)=-1} W_\alpha + \frac{b-m+r_2-1}{2}(W'+W'')$
$q \equiv 7 \pmod 8$	$+ (1-r_2)\sum_{\beta(i)=-1} X_\beta + \frac{b-m}{2}(X'+X'')$
$q \equiv 1 \pmod 3$	$+ (r_3-1)\sum_{\alpha(\omega)\neq 1} W_\alpha$
$q \equiv 2 \pmod 3$	$+ (1-r_3)\sum_{\beta(\omega)\neq 1} X_\beta$
$q \equiv 1 \pmod 7$	$+ \sum_{k=1}^{r_7} \sum_{\alpha(\phi)=e^{\pm\frac{2\pi i k}{7}}} W_\alpha - \sum_{alpha(\phi)\neq 1} W_\alpha$
$q \equiv 6 \pmod 7$	$+ \sum_{\beta(\phi)\neq 1} X_\beta - \sum_{k=1}^{r_7} \sum_{\beta(\phi)=e^{\pm\frac{2\pi i k}{7}}} X_\beta$

4.4. **Action on holomorphic differentials.** As a corollary, it is an easy exercise now to compute explicitly the decomposition

$$H^1(X,\mathbb{C}) = H^0(X,\Omega^1) \oplus \overline{H^0(X,\Omega^1)} = L(K_X) \oplus \overline{L(K_X)},$$

into irreducible G-modules, where K_X is a canonical divisor of X. The action of G on the complex conjugate vector space $\overline{L(K_X)}$ of $L(K_X)$ will be by the complex conjugate (contragredient) representation. The Riemann-Hurwitz theorem tells us that

$$\begin{aligned} K_X &= \pi^*(K_{\mathbb{P}^1}) + R \\ &= -2D_1 + D_2 + 2D_3 + 6D_7 \end{aligned}$$

where R is the ramification divisor. Thus the equivariant degree of K_X is $\deg_{eq}(K_X) = -2\cdot\mathbb{C}[G] + \deg_{eq}(R)$. Note from the preliminary equivariant degree calculations,

that

$$\begin{aligned}\deg_{eq}(R) &= \deg_{eq} D_2 + \deg_{eq} 2D_3 + \deg_{eq} 6D_7 \\ &= \mathrm{Ind}_{H_2}^G \theta_2 + 2\,\mathrm{Ind}_{H_3}^G \theta_3 + \sum_{k=1}^{6} \mathrm{Ind}_{H_7}^G \theta_7^k \\ &= 2\Gamma_{H_2} + 2\Gamma_{H_3} + 2\Gamma_{H_7} \\ &= 2\tilde{\Gamma}.\end{aligned}$$

Therefore, using the equivariant Riemann-Roch formula (1),

$$L(K_X) = \tilde{\Gamma} - \mathbb{C}[G]. \tag{8}$$

We will see in the next section that this is invariant under complex conjugation, so that as G-modules, $H^1(X, \mathbb{C}) \cong 2L(K_X)$.

Using the results of section 4.1, we obtain the following.

Theorem 5. *The G-module structure of $L(K) = H^0(X, \Omega^1)$ is as follows:*

- *If $q \equiv 1$, 97, or 113 (mod 168), then*

$$L(K_X) = \frac{\lfloor \frac{q}{84} \rfloor - 1}{2}(W' + W'') + \sum_{\beta} \left(\lfloor \frac{q}{84} \rfloor + 1 - N_\beta \right) X_\beta + \lfloor \frac{q}{84} \rfloor V + \sum_{\alpha} \left(\lfloor \frac{q}{84} \rfloor - 1 + N_\alpha \right) W_\alpha.$$

- *If $q \equiv 43$ (mod 168), then*

$$L(K_X) = \lfloor \frac{q}{84} \rfloor \left[\frac{1}{2}(X' + X'') + V \right] + \sum_{\beta} \left(\lfloor \frac{q}{84} \rfloor + 1 - N_\beta \right) X_\beta + \sum_{\alpha} \left(\lfloor \frac{q}{84} \rfloor - 1 + N_\alpha \right) W_\alpha.$$

- *If $q \equiv 13$, 29, or 85 (mod 168), then*

$$L(K_X) = \lfloor \frac{q}{84} \rfloor \left[\frac{1}{2}(W' + W'') + V \right] + \sum_{\beta} \left(\lfloor \frac{q}{84} \rfloor + 1 - N_\beta \right) X_\beta + \sum_{\alpha} \left(\lfloor \frac{q}{84} \rfloor - 1 + N_\alpha \right) W_\alpha.$$

- *If $q \equiv 127$ (mod 168), then*

$$L(K_X) = \frac{\lfloor \frac{q}{84} \rfloor + 1}{2}(X' + X'') + \sum_{\beta} \left(\lfloor \frac{q}{84} \rfloor + 1 - N_\beta \right) X_\beta + \lfloor \frac{q}{84} \rfloor V + \sum_{\alpha} \left(\lfloor \frac{q}{84} \rfloor - 1 + N_\alpha \right) W_\alpha.$$

- *If $q \equiv 41$ (mod 168), then*

$$L(K_X) = \frac{\lceil \frac{q}{84} \rceil - 1}{2}(W' + W'') + \sum_{\beta} \left(\lceil \frac{q}{84} \rceil + 1 - N_\beta \right) X_\beta + \lceil \frac{q}{84} \rceil V + \sum_{\alpha} \left(\lceil \frac{q}{84} \rceil - 1 + N_\alpha \right) W_\alpha.$$

- *If $q \equiv 83$, 139, or 155 (mod 168), then*

$$L(K_X) = \lceil \frac{q}{84} \rceil \left[\frac{1}{2}(X' + X'') + V \right] + \sum_{\beta} \left(\lceil \frac{q}{84} \rceil + 1 - N_\beta \right) X_\beta + \sum_{\alpha} \left(\lceil \frac{q}{84} \rceil - 1 + N_\alpha \right) W_\alpha.$$

- *If $q \equiv 125$ (mod 168), then*

$$L(K_X) = \lceil \frac{q}{84} \rceil \left[\frac{1}{2}(W' + W'') + V \right] + \sum_{\beta} \left(\lceil \frac{q}{84} \rceil + 1 - N_\beta \right) X_\beta + \sum_{\alpha} \left(\lceil \frac{q}{84} \rceil - 1 + N_\alpha \right) W_\alpha.$$

- *If* $q \equiv 55,\ 71,$ *or* $167 \pmod{168}$, *then*

$$L(K_X) = \frac{\lceil \frac{q}{84} \rceil + 1}{2}(X' + X'') + \sum_{\beta} \left(\lceil \frac{q}{84} \rceil + 1 - N_\beta \right) X_\beta + \lceil \frac{q}{84} \rceil V + \sum_{\alpha} \left(\lceil \frac{q}{84} \rceil - 1 + N_\alpha \right) W_\alpha.$$

5. Galois action

As discussed in section 3, there is a Galois action on the set of equivalence classes of irreducible representations of $PSL(2,q)$. One question of obvious interest is whether the modules we have computed are invariant under this action.

Theorem 6. *The ramification module is Galois-invariant.*

Proof. Recall from section 3 that the Galois group $\mathcal{G}$ permutes mth roots of unity, where $m = q(q^2-1)/4$. It acts on representations of $PSL(2,q)$ by permuting character values. Thus it fixes the trivial representation and the q-dimensional representation V, whose character values are rational. It will act as a permutation on the representations W_α and on the representations X_β. Lastly, it will act as an involution on either the representations W' and W'' or X' and X''.

Because the multiplicities of W' and W'' or X' and X'' are the same in the ramification module, the Galois action will be invariant on this component. The multiplicity of a representation W_α or X_β in the ramification module depends on the number N_α or N_β, which is determined by the value of the character α or β on the special numbers i, ω, and ϕ. In fact, the numbers N_α and N_β are determined only by whether these character values are equal to 1 or not equal to 1. Since an element of the Galois group will take a character value to a power of itself, the Galois action must preserve the numbers N_α and N_β. Therefore this component of the ramification module is invariant as well.

□

Since the ramification module is Galois-invariant, and of course the regular representation is Galois-invariant, $L(K_X)$ will be Galois invariant. In particular, as stated in section 4.4, $L(K_X)$ will be invariant under complex conjugation. For a general divisor D, the Riemann-Roch space $L(D)$ will be Galois-invariant if and only if the equivariant degree of D is.

Theorem 7. *Let* $D = r_1D_1 + r_2D_2 + r_3D_3 + r_7D_7$ *be a* G*-invariant divisor. Then the equivariant degree of* D *is Galois-invariant if* $r_7 \in \{0,3,6\} \pmod 7$.

Proof. As in section 4.2, multiples of 2 in r_2, 3 in r_3, and 7 in r_7 can be absorbed into the r_1D_1 term without affecting the equivariant degree. Therefore we may assume that $0 \le r_2 \le 1$, $0 \le r_3 \le 2$, and $0 \le r_7 \le 6$.

The result can again be seen by looking at the multiplicities of representations permuted by the Galois group. The multiplicities of W' and W'' or X' and X'' are the same. By (7), the multiplicity of a representation W_α or X_β depends on r_2, r_3, and r_7, and not on r_1. Again, the Galois action will not permute a representation W_α with $\alpha(i) = 1$ with one with $\alpha(i) \neq 1$; similarly for X_β, and for ω. However, it could permute for example a representation W_α with $\alpha(\phi) = e^{\frac{2\pi i}{7}}$ with one with $\alpha(\phi) = e^{\frac{4\pi i}{7}}$. Thus the equivariant degree may not be Galois-invariant unless the multiplicities of these representations are equal. In the cases where $r_7 \in \{0,3,6\}$, then these multiplicities will be equal; otherwise they will not.

□

Note that for some values of q, the equivariant degree may be Galois-invariant even if r_7 is not 0, 3, or 6.

A previous result of the first two authors (see [JK1]) gives a simpler formula (see equation 2) to compute the multiplicity of an irreducible representation in the ramification module, when the ramification module is Galois-invariant. In the example at hand, if $r_7 \in \{0,3,6\}$, then since the equivariant degree is a multiple of the H_7 component of the ramification module, a slight modification of this formula gives an easy computation of the equivariant degree and therefore the Riemann-Roch space.

Corollary 8. *Let $D = r_1 D_1 + r_2 D_2 + r_3 D_3 + r_7 D_7$, with $0 \leq r_2 \leq 1$, $0 \leq r_3 \leq 2$, and $r_7 \in \{0,3,6\}$. Then*

$$\begin{aligned} L(D) &= \bigoplus_{\pi \in G^*} \Big[(1 + r_1 + r_2 + \frac{r_3}{2} + \frac{r_7}{6}) \dim \pi \\ &\quad + (\frac{1}{2} - r_2) \dim \pi^{H_2} + (\frac{1}{2} - \frac{r_3}{2}) \dim \pi^{H_3} + (\frac{1}{2} - \frac{r_7}{6}) \dim \pi^{H_7} \Big] \pi. \end{aligned}$$

Note that in spite of appearances, the multiplicity of each irreducible representation will in fact be an integer.

Proof. We see from the calculations in section 4.2 that the equivariant degree of D is equal to

$$\begin{aligned} \deg_{eq}(D) &= r_1 \mathbb{C}[G] + 2 r_2 \Gamma_{H_2} + r_3 \Gamma_{H_3} + \frac{r_7}{3} \Gamma_{H_7} \\ &= \bigoplus_{\pi \in G^*} \Big[(r_1 + r_2 + \frac{r_3}{2} + \frac{r_7}{6}) \dim \pi \\ &\quad - r_2 \dim \pi^{H_2} - \frac{r_3}{2} \dim \pi^{H_3} - \frac{r_7}{6} \dim \pi^{H_7} \Big] \pi. \end{aligned}$$

The ramification module is

$$\tilde{\Gamma}_G = \bigoplus_{\pi \in G^*} \left[\sum_{\ell \in 2,3,7} (\dim \pi - \dim (\pi^{H_\ell})) \frac{1}{2} \right] \pi.$$

This sum splits into $\tilde{\Gamma}_G = \Gamma_{H_2} + \Gamma_{H_3} + \Gamma_{H_7}$ in the obvious way along the inner sum. Putting these together using the equivariant Riemann-Roch formula (1), we obtain the desired result.

□

References

[B] N. Borne, "Une formule de Riemann-Roch equivariante pour des courbes," Can. J. Math. **55**(2003)693-710. (see also thesis, Univ. Bordeaux, 1999. Available on the web at `http://www.dm.unibo.it/~borne/`)

[C] M. Conder, "Hurwitz groups: a brief survey," Bull. Amer. Math. Soc. 23(2) (1990) 359–370.

[EL] G. Ellingsrud and K. Lønsted, "An equivariant Lefschetz trace formula for finite reductive groups," Math Ann **251**(1980), 253-261.

[FH] W. Fulton and J. Harris, **Representation theory: a first course**, Springer-Verlag, 1991.

[GAP] The GAP Group, **GAP – Groups, Algorithms, and Programming, Version 4.4**; 2002, (`http://www.gap-system.org`).

[Ja] G. Janusz, "Simple components of $\mathbb{Q}[SL(2,q)]$," Comm. Alg. 1(1974)1-22.

[JK1] D. Joyner and A. Ksir, "Decomposing representations of finite groups on Riemann-Roch spaces," to appear in *Proceedings of the American Mathematical Society*, math.AG/0312383.

[JK2] ——, "Modular representations on some Riemann-Roch spaces of modular curves $X(N)$," in **Computational Aspects of Algebraic Curves**, (Editor: T. Shaska) Lecture Notes in Computing, WorldScientific, 2005, math.AG/0502586.

[K] E. Kani, "The Galois-module structure of the space of holomorphic differentials of a curve," J. Reine Angew. Math. **367** (1986), 187-206.

[L] M. Larsen, "How often is $84(g-1)$ achieved?" Israel J. Math. 126 (2001) 1–16.

[M] A.M. Macbeath, "On a theorem of Hurwitz," Proc. Glasgow Math. Assoc. 5 (1961) 90–96.

[N] S. Nakajima, "Galois module structure of cohomology groups for tamely ramified coverings of algebraic varieties," J. Number Theory 22 (1986) 115-123.

Mathematics Department, US Naval Academy, Annapolis, MD, 21402
E-mail address: wdj@usna.edu

Mathematics Department, US Naval Academy, Annapolis, MD, 21402
E-mail address: ksir@usna.edu

Mathematics Department, Ohio State University, Columbus, OH, 43210.
E-mail address: vogeler@math.ohio-state.edu

ALBANIAN JOURNAL
OF MATHEMATICS
Volume 1, Number 2, Pages 87–98
ISSN 1930-1235: (electronic version)

REFLECTION AND TRANSMISSION OF WAVES AT AN ELASTIC INTERFACE OF TWO HALF SPACES SUBJECT TO PURE SHEAR

WASIQ HUSSAIN

Abstract. We study the effect of *pure shear* on the reflection and transmission of plane waves at the boundary between two half-spaces of incompressible isotropic elastic material. The half-spaces consist of the same material and are subjected to pure shear deformation with their principal axes aligned. The objective is to highlight the dependence of the amplitudes of the elastic waves on the finite pure shear deformation and thereby to provide a theoretical framework for the non-destructive evaluation at the shear interface.

When the first half-space corresponds to a certain class of constitutive laws and the second half-space *(not in this special class),* depending upon the angle of incidence, the material properties, and the magnitudes of deformations, it is shown that a homogeneous plane (SV) wave propagating in the plane of pure shear gives rise to a reflected wave (with angle of reflection equal to the angle of incidence) together with an interfacial wave in the same half-space, while in the other half-space there is a transmitted wave accompanied by an interfacial wave.

The dependence of the amplitudes of the reflected, transmitted, and interfacial waves on the angle of incidence and the states of deformation is illustrated graphically. The results described here provide a basis for the characterization of material properties and the finite homogeneous shear deformation.

1. Introduction

In [1] Hussain and Ogden have examined the effect of a finite simple shear deformation on the reflection of superimposed infinitesimal plane waves incident on the boundary of a half-space of incompressible isotropic elastic material. References to the literature concerned with reflection at the boundary of a finitely deformed half-space are contained in [2,3].

In the present paper the effect of *pure shear* on the reflection and *transmission* of plane (shear) waves at the boundary between two half-spaces which consist of the same material (but with *different* strain-energy functions) is considered. This problem of *mixed* strain-energy functions has *not* apparently been considered previously. The configuration is intended to describe the finite pure homogeneous shear deformation associated with the two half-spaces with a view to study theoretically the vibration and wave propagation characteristics of rubber like solids.

Received by the editors December 11, 2006 and, in revised form, March 30, 2007.
2000 *Mathematics Subject Classification.* 74Bxx; 74Jxx.
Key words and phrases. Elastic waves, pure shear, reflection, transmission, non-linear elasticity.

The required equations and notations are summarized in Section 2. In Section 3 the propagation of plane harmonic waves is discussed with reference to the *slowness curves* appropriate for the two distinct classes of strain-energy functions.

The amplitudes of the reflected, transmitted and interfacial waves are calculated in Section 4 when a given homogeneous plane (shear) wave is incident on the boundary. *A combined case* of (distinct) strain-energy functions is discussed. In the paper by Dey and Addy [4] reflection and refraction of plane waves at an interface is discussed, which, as pointed out by Norris [5], contains fundamental errors.

For each angle of incidence a single reflected wave, with angle of reflection equal to the angle of incidence, is generated when a homogeneous plane (SV) wave is incident on the boundary from one half-space, and it is accompanied by an interfacial wave. In $x_2 > 0$ a transmitted wave and an interfacial wave are generated for *all* angles of incidence.

The theory in Section 4 is illustrated in Section 5 using graphical results to show the dependence of the amplitudes of the waves on the angle of incidence for representative values of the deformation parameters.

Finally, in Section 6 conclusions are given, describing the significance of the results obtained along with the research options for the extension of the analysis done in this paper.

2. Basic equations

We identify the undeformed configuration of the material, $\mathcal{B}_0$ say, and let a material particle in $\mathcal{B}_0$ be labelled by its three dimensional position vector $\mathbf{X}$. Let $\mathbf{x}$ be the position vector of the same particle in the deformed configuration, $\mathcal{B}$ say. We write the deformation of the material from $\mathcal{B}_0$ to $\mathcal{B}$, $\boldsymbol{\chi}$ say, as

$$\mathbf{x} = \boldsymbol{\chi}(\mathbf{X}), \quad \mathbf{X} \in \mathcal{B}_0.$$

The deformation gradient tensor $\mathbf{A}$ is defined as

$$\mathbf{A} = \mathrm{Grad}\boldsymbol{\chi},$$

where Grad denotes the gradient with respect to $\mathbf{X}$, and is subject to the usual condition

$$\det \mathbf{A} > 0.$$

The polar decomposition theorem enables the second order tensor $\mathbf{A}$ to be written as

$$\mathbf{A} = \mathbf{VR},$$

where $\mathbf{R}$ is a proper orthogonal tensor and $\mathbf{V}$ is the symmetric and positive definite *left stretch tensor*.

Let $\mathbf{dX}$ is an arbitrary line element based at $\mathbf{X}$ in the reference configuration and $\mathbf{dx}$ is the corresponding line element at $\mathbf{x}$ in the deformed configuration. The *stretch* in the direction of $\mathbf{dX}$ at $\mathbf{X}$, is defined as the ratio of current to reference lengths of a line element and is given by

$$\frac{|\mathbf{dx}|}{|\mathbf{dX}|} = \lambda(\mathbf{M}),$$

where $\mathbf{M}$ is a unit vector along $\mathbf{dX}$.

The volume elements dV and dv in the reference and deformed configurations respectively are related by

$$dv = (det\mathbf{A})\, dV,$$

therefore for a volume preserving deformation we have

$$\det \mathbf{A} \equiv \lambda_1\lambda_2\lambda_3 = 1, \tag{1}$$

where $\lambda_i\,(>0)\,(i=1,2,3)$ are the eigenvalues (*principal stretches*), corresponding to the eigenvectors $\mathbf{v}_i\,(i=1,2,3)$, of the symmetric and positive definite tensor $\mathbf{V}$.

Let $\mathbf{S}$ denote the nominal stress tensor. Then, the equilibrium equation, in the absence of body forces, is

$$\mathrm{Div}\,\mathbf{S} = \mathbf{0},$$

where Div is the divergence operator in the reference configuration and $\mathbf{0} \in R^3$.

The measure of the energy stored per unit reference volume in the material as a result of deformation is called the elastic *stored energy function*. More commonly, the phrase *strain-energy function* is used to describe W (say) and this is the terminology we adopt in this paper.

For a (homogeneous) elastic material with strain-energy function $W = W(\mathbf{A})$ per unit volume, subject to the incompressibility constraint given by Eq. (1), we have

$$\mathbf{S} = \frac{\partial W}{\partial \mathbf{A}} - p\mathbf{A}^{-\mathbf{1}},$$

where p is a Lagrange multiplier, which can be identified as a *hydrostatic pressure* associated with the incompressibility constraint. In general, it is a scalar function of time t, and in the literature is often introduced with the opposite sign.

An *isotropic elastic material* is an elastic material whose symmetry group contains the proper orthogonal group for at least one reference configuration. In such a reference configuration the mechanical response of the material exhibits no preferred direction, and it is this property that characterizes isotropy. If the material is isotropic, W depends symmetrically on $\lambda_1, \lambda_2, \lambda_3$ subject to Eq. (1) and we write $W(\lambda_1, \lambda_2, \lambda_3)$.

For the isotropic material, the principal Cauchy stresses are given by

$$\sigma_i = \lambda_i \frac{\partial W}{\partial \lambda_i} - p, \qquad i \in \{1,2,3\}.$$

For (plane strain) deformations confined to the $(1,2)$-plane, we may set $\lambda_3 = 1$, so that Eq. (1) reduces to

$$\lambda_1\lambda_2 = 1.$$

Homogeneous *pure shear* deformation is defined by

$$\lambda_1 = \lambda \neq 1, \;\; \lambda_2 = \lambda^{-1}, \; \lambda_3 = 1 \text{ with } \sigma_1 \neq 0, \; \sigma_2 = 0,$$

where a non-vanishing stress σ_3 is required to maintain $\lambda_3 = 1$. Superimposed on the deformation just described we consider incremental motions in the (x_1, x_2)-plane with displacement vector $\mathbf{v}$ having components

$$v_1(x_1, x_2, t), \quad v_2(x_1, x_2, t), \quad v_3 = 0.$$

The (linearized) incremental incompressibility condition $\operatorname{div}\mathbf{v}=0$ enables v_1, v_2 to be expressed in terms of a scalar function, $\psi(x_1,x_2,t)$ say, so that

$$v_1=\psi_{,2}, \qquad v_2=-\psi_{,1}, \tag{2}$$

where $,i$ denotes $\partial/\partial x_i$, $i\in\{1,2\}$.

The incremental nominal stress tensor is denoted by $\mathbf{\Sigma}$ when referred to the deformed configuration. Its components are given by

$$\Sigma_{ji}=\mathcal{A}_{0jilk}v_{k,l}+pv_{j,i}-\pi\delta_{ij}, \tag{3}$$

where π is the increment in p and $\mathcal{A}_{0jilk}$ are the components of the fourth-order tensor $\boldsymbol{\mathcal{A}_0}$ of instantaneous elastic moduli (see, for example, Ogden [6]).

The components of $\boldsymbol{\mathcal{A}_0}$ in terms of the derivatives of the strain-energy function W are given by

$$\begin{aligned}
&\mathcal{A}_{0iijj}=\lambda_i\lambda_jW_{ij},\\
&\mathcal{A}_{0ijij}=\frac{(\lambda_iW_i-\lambda_jW_j)\lambda_i^2}{(\lambda_i^2-\lambda_j^2)} \qquad i\neq j,\quad \lambda_i\neq\lambda_j,\\
&\mathcal{A}_{0ijij}=\frac{1}{2}(\mathcal{A}_{0iiii}-\mathcal{A}_{0iijj}+\lambda_iW_i) \qquad i\neq j,\quad \lambda_i=\lambda_j,\\
&\mathcal{A}_{0ijji}=\mathcal{A}_{0jiij}=\mathcal{A}_{0ijij}-\lambda_iW_i \qquad i\neq j,
\end{aligned} \tag{4}$$

where $W_i=\partial W/\partial\lambda_i$, $W_{ij}=\partial^2W/\partial\lambda_i\partial\lambda_j$ and there is no summation over repeated indices. Here, the components $\mathcal{A}_{0jilk}$ are constants because the deformation under consideration is homogeneous.

The equation of motion is given by

$$\mathcal{A}_{0jilk}v_{k,jl}-\pi_{,i}=\rho\ddot{v}_i, \qquad i\in\{1,2\}, \tag{5}$$

where ρ is the mass density of the material and there is summation from 1 to 2 over repeated indices. The equations of motion given by Eq. (5) yield, on restriction to the considered plane motion,

$$\begin{aligned}
&(\mathcal{A}_{01111}-\mathcal{A}_{01122}+p)v_{1,11}-\pi_{,1}+\mathcal{A}_{02121}v_{1,22}+(\mathcal{A}_{02121}-\sigma_2)v_{2,12}=\rho\ddot{v}_1,\\
&(\mathcal{A}_{02222}-\mathcal{A}_{02211}+p)v_{2,22}-\pi_{,2}+\mathcal{A}_{01212}v_{2,11}+(\mathcal{A}_{02121}-\sigma_2)v_{1,12}=\rho\ddot{v}_2,
\end{aligned} \tag{6}$$

where a superposed dot indicates the material time derivative.

Elimination of π from Eq. (6), and use of Eq. (2) yields an equation for ψ, namely

$$\alpha\psi_{,1111}+2\beta\psi_{,1122}+\gamma\psi_{,2222}=\rho(\ddot{\psi}_{,11}+\ddot{\psi}_{,22}), \tag{7}$$

as given in [1], where the constants α,β,γ are defined by

$$\alpha=\mathcal{A}_{01212}, \quad \gamma=\mathcal{A}_{02121}, \quad 2\beta=\mathcal{A}_{01111}+\mathcal{A}_{02222}-2\mathcal{A}_{01122}-2\mathcal{A}_{01221}. \tag{8}$$

From Eq. (3), by using Eq. (2), the shear and normal components of the incremental nominal traction Σ_{21}, Σ_{22} on a plane $x_2=$ constant are expressible in terms of ψ through

$$\begin{aligned}
\Sigma_{21}&=\gamma\psi_{,22}-(\gamma-\sigma_2)\psi_{,11},\\
-\Sigma_{22,1}&=(2\beta+\gamma-\sigma_2)\psi_{,112}+\gamma\psi_{,222}-\rho\ddot{\psi}_{,2},
\end{aligned} \tag{9}$$

in the latter of which the incremental hydrostatic pressure π has been eliminated by differentiating Σ_{22} with respect to x_1 and then using first equation in Eq. (6).

For any type of detail discussion, related to basic equations, the reader is requested to see Ogden [6].

3. Plane waves

We consider time-harmonic homogeneous plane waves of the form

$$\psi = A\exp[ik(x_1\cos\theta + x_2\sin\theta - ct)], \tag{10}$$

where A is a constant, c (>0) the wave speed, k (>0) the wave number and $(\cos\theta, \sin\theta)$ the direction cosines of the direction of propagation of the wave in the (x_1, x_2)-plane. Substitution of Eq. (10) into Eq. (7) gives

$$\alpha\cos^4\theta + 2\beta\sin^2\theta\cos^2\theta + \gamma\sin^4\theta = \rho c^2. \tag{11}$$

Equation (11) is a relationship between the wave speed and the propagation direction in the (x_1, x_2)-plane and is called the *propagation condition.* The material constants are taken to satisfy the strong ellipticity inequalities

$$\alpha > 0, \quad \gamma > 0, \quad \beta > -\sqrt{\alpha\gamma}, \tag{12}$$

and it is clear from Eq. (11) that $\rho c^2 > 0$ if and only if Eq. (12) hold.

Similarly, from Eq. (7), for an inhomogeneous plane wave of the form

$$\psi = \hat{A}\exp[ik'(x_1 - imx_2 - c't)], \tag{13}$$

we obtain

$$\alpha - 2\beta m^2 + \gamma m^4 = \rho(1-m^2)c'^2, \tag{14}$$

which relates the wave speed c' to the 'inhomogeneity factor' m. Note that the wave decays exponentially as $x_2 \to -\infty(+\infty)$ provided m has positive (negative) real part.

We now consider two half-spaces of the same incompressible isotropic elastic material. The half-spaces are subjected to pure shear deformation and then bonded along their common (plane) boundary in such a way that the principal directions of strain are aligned, one direction being normal to the interface.

Let λ_1, λ_2, λ_3 be the stretches associated with the half-spaces $x_2 < 0$, $x_2 > 0$, with strain energy function W and the material constants α, β, γ defined by Eq. (4) with Eq. (8).

We take the deformation to correspond to pure shear with $\lambda_3 = 1$ so that, with reference to the incompressibility condition (1), we introduce the notation λ such that

$$\lambda_1 = \lambda_2^{-1} = \lambda.$$

We consider two distinct cases corresponding to different strain-energy functions. For these either $2\beta = \alpha + \gamma$ or $2\beta \neq \alpha + \gamma$.

3.1. **Case A:** $2\beta = \alpha + \gamma$.

For this case equations Eq. (11) and Eq. (14) reduce to

$$\alpha\cos^2\theta + \gamma\sin^2\theta = \rho c^2 \tag{15}$$

and

$$(m^2-1)(\alpha - \gamma m^2 - \rho c'^2) = 0 \tag{16}$$

respectively.

In terms of the *slowness vector* (s_1, s_2) defined by

$$(s_1, s_2) = (\cos\theta, \sin\theta)/c$$

Eq. (15) becomes the *slowness* curve

$$\lambda^4 {s_1}^2 + {s_2}^2 = \overline{\rho}, \tag{17}$$

in the (s_1, s_2)-space, where $\overline{\rho}$ is defined by

$$\overline{\rho} = \rho/\gamma, \tag{18}$$

and $\alpha/\gamma = \lambda^4$ follows from Eq. (4) and Eq. (8).

By using the dimensionless notation $(\overline{s}_1, \overline{s}_2)$ defined by

$$(\overline{s}_1, \overline{s}_2) \equiv (s_1, s_2)/\sqrt{\overline{\rho}}, \tag{19}$$

we can write Eq. (17) as

$$\lambda^4 {\overline{s}_1}^2 + {\overline{s}_2}^2 = 1. \tag{20}$$

3.2. **Case B:** $2\beta \neq \alpha + \gamma$**.** In this case we take the strain-energy function to satisfy $\beta = \sqrt{\alpha\gamma}$ which was used by Hussain and Ogden in [1]. Then Eq. (11) takes the form

$$[\sqrt{\alpha}\cos^2\theta + \sqrt{\gamma}\sin^2\theta]^2 = \rho c^2 \tag{21}$$

and Eq. (14) becomes

$$(\sqrt{\alpha} - \sqrt{\gamma}m^2)^2 = \rho(1 - m^2){c'}^2. \tag{22}$$

The slowness curve corresponding to Eq. (21) is given by

$$[\lambda^2\overline{s}_1^2 + \overline{s}_2^2]^2 = {\overline{s}_1}^2 + {\overline{s}_2}^2, \tag{23}$$

in dimensionless form with the notation given by Eq. (19) and $\overline{\rho}$ defined by Eq. (18). We now show graphically the dependence of the slowness curves on λ for both classes of strain-energy functions in $(\overline{s}_1, \overline{s}_2)$-space with reference to Eq. (20) and Eq. (23)(See Fig. 1-2).

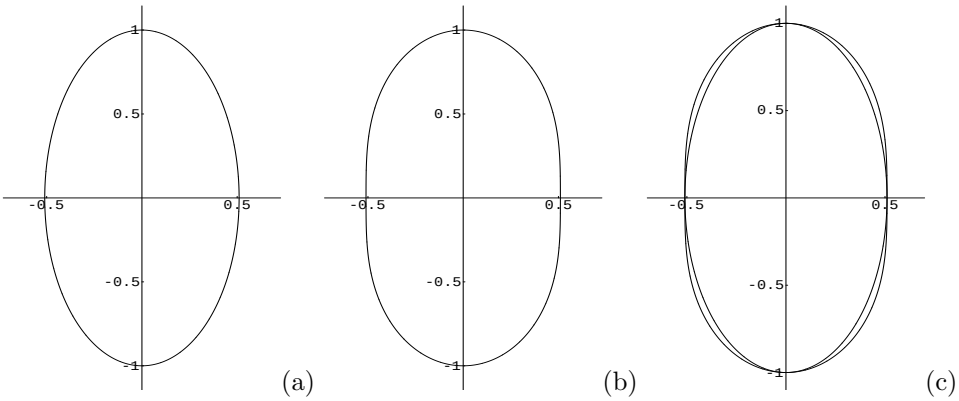


FIGURE 1. Slowness curves in $(\overline{s}_1, \overline{s}_2)$-space for $\lambda = 1.4$ with (a) $2\beta = \alpha + \gamma$, (b) $2\beta \neq \alpha + \gamma$, (c) the superposition of Figs. in (a) and (b).

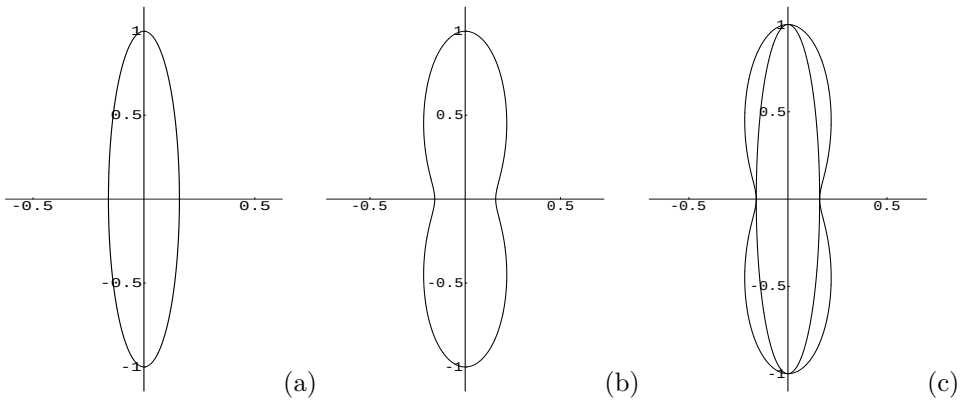


FIGURE 2. Slowness curves in $(\bar{s}_1, \bar{s}_2)$-space for $\lambda = 2.5$ with (a) $2\beta = \alpha + \gamma$, (b) $2\beta \neq \alpha + \gamma$, (c) the superposition of Figs. in (a) and (b).

4. Reflection and transmission at the interface

The boundary conditions corresponding to continuous displacement are $v_1 = v_1^*,\ v_2 = v_2^*$ on $x_2 = 0$, where v_1, v_2 are the displacement components in $x_2 < 0$ and v_1^*, v_2^* are those in $x_2 > 0$. From Eq. (2) these boundary conditions can be written in terms of the scalar functions ψ and ψ^* as

$$\psi_{,1} = \psi^*_{,1},\ \psi_{,2} = \psi^*_{,2} \quad \text{on} \quad x_2 = 0, \tag{24}$$

where ψ^* is the counterpart of ψ for $x_2 > 0$.

The boundary conditions for continuous incremental traction on the interface are

$$\Sigma_{21} = \Sigma^*_{21}, \quad \Sigma_{22} = \Sigma^*_{22} \quad \text{on} \quad x_2 = 0, \tag{25}$$

where $\Sigma_{21},\ \Sigma_{22}$ are the traction components in $x_2 < 0$ and $\Sigma^*_{21},\ \Sigma^*_{22}$ are those in $x_2 > 0$.

From Eq. (9) the boundary conditions given by Eq. (25) take the forms

$$\begin{aligned} \psi_{,11} - \psi_{,22} &= \psi^*_{,11} - \psi^*_{,22}, \\ (2\beta + \gamma)(\psi_{,112} - \psi^*_{,112}) + \gamma(\psi_{,222} - \psi^*_{,222}) - \rho(\ddot{\psi}_{,2} - \ddot{\psi}^*_{,2}) &= 0, \end{aligned} \tag{26}$$

in terms of ψ and ψ^*, where, in order to obtain the second equation in Eq. (26), the second equation in Eq. (25) has been replaced by $\Sigma_{22,1} = \Sigma^*_{22,1}$ and use made of the second equation in Eq. (9) and its counterpart for $x_2 > 0$.

We now consider a wave incident on the boundary $x_2 = 0$ from the region $x_2 < 0$ with direction of propagation $(\cos\theta, \sin\theta)$ in the (x_1, x_2)- plane and speed c. Because of the symmetry of slowness curves with respect to the normal direction to the interface we henceforth, without loss of generality, restrict attention to values of θ in the interval $[0, \pi/2]$. We write the solution comprising the incident wave, a reflected wave (with angle of reflection equal to the angle of incidence) and an interfacial wave in $x_2 < 0$ as

$$\begin{aligned} \psi = A\exp[ik(x_1\cos\theta + x_2\sin\theta - ct)] + AR\exp[ik(x_1\cos\theta - x_2\sin\theta - ct)] \\ + AR'\exp[ik'(x_1 - imx_2 - c't)], \end{aligned} \tag{27}$$

where R is the reflection coefficient and R' measures the amplitude of the interfacial wave. The notations k', m, c' are as used in Eq. (13) and m has positive real part.

In the half-space $x_2 > 0$ we write the solution comprising a transmitted and an interfacial wave in the form

(28)
$$\psi^* = AR^* \exp[ik^*(x_1 \cos\theta^* + x_2 \sin\theta^* - c^* t)] + AR^{*'} \exp[ik^{*'}(x_1 + im^* x_2 - c^{*'} t)],$$

where R^* is the transmission coefficient and $R^{*'}$ is the analogue of R' for $x_2 > 0$. The transmitted wave has direction of propagation $(\cos\theta^*, \sin\theta^*)$, wave number k^* and speed c^*, while $k^{*'}$, m^*, $c^{*'}$ are the counterparts of k', m, c'. Note that the interfacial wave decays as $x_2 \to \infty$ provided m^* has positive real part.

According to the Snell's law we have

$$\cos\theta/c = 1/c' = \cos\theta^*/c^* = 1/c^{*'}. \tag{29}$$

Eq. (29) states in particular, that the first components of the slowness vectors for each homogeneous plane wave interacting at the boundary $x_2 = 0$ are equal.

Thus, by reference to the slowness curves (superimposed) as exemplified in Fig. 1(c) and Fig. 2(c), the range of angles of incidence for which a transmitted wave exists can be identified. In Figs. 1(c) and 2(c), for example, if the *inner curve* corresponds to $x_2 < 0$ there is, for every angle of incidence (i.e. for every s_1 associated with the curve) a point on the outer curve (corresponding to $x_2 > 0$), and hence a transmitted wave.

We now examine here the case in which $2\beta = \alpha + \gamma\,(x_2 < 0)$, $2\beta \neq \alpha + \gamma\,(x_2 > 0)$. Analogous results, obtainable for $2\beta \neq \alpha + \gamma\,(x_2 < 0)$, $2\beta = \alpha + \gamma\,(x_2 > 0)$ will be discussed elsewhere.

4.1. $2\beta = \alpha + \gamma\,(x_2 < 0)$, $2\beta \neq \alpha + \gamma\,(x_2 > 0)$**.** In this case we see from Eq. (16) that $m = \pm 1$, which yields an interfacial wave in the half-space $x_2 < 0$ for $m = 1$. The zeros of the other quadratic factor in Eq. (16) correspond to $m = i\tan\theta$ and $m = -i\tan\theta$ which are associated, respectively, with the incident and reflected waves in $x_2 < 0$.

In $x_2 > 0$, from the counterpart of Eq. (22), after using $\alpha/\gamma = \lambda^4$ and Snell's law $\cos\theta^*/c^* = 1/c^{*'}$, we have

$$(m^{*2} + t^{*2})[m^{*2}(1 + t^{*2}) - t^{*2} + \lambda^2(\lambda^2 - 2)] = 0,$$

where $t^* = \tan\theta^*$.

The solution $m^* = it^*$ corresponds to a transmitted wave provided t^* is real and positive. The other relevant solution is

$$m^* = \pm\sqrt{1 - (\lambda^2 - 1)^2/(1 + t^{*2})} \tag{30}$$

with the plus sign when m^* is real. The nature of m^* in Eq. (30) depends on that of t^*, which is obtained by using the propagation condition given by Eq. (15) and the counterpart of Eq. (21) for the (transmitted) wave with direction of propagation $(\cos\theta^*, \sin\theta^*)$ and speed c^* together with Snell's law Eq. (29). This gives a quadratic for t^{*2}, which we write as

$$t^{*4} + t^{*2}(2\lambda^2 - \lambda^4 - t^2) - t^2 = 0, \tag{31}$$

and the notation $t = \tan\theta$ has been introduced. Note that t should be distinguished from the time variable t used earlier.

If t_1^{*2} and t_2^{*2} are the roots of Eq. (31) then we have

$$t_1^{*2} t_2^{*2} = -t^2,$$

which shows that there is one positive and one negative solution for t^{*2}, and hence one transmitted and one interfacial wave. See also Fig. 1(c) and Fig. 2(c).

When there is no refraction, i.e. a transmitted wave has the same direction of propagation as the incident wave ($\theta^* = \theta$). For this to be the case we must have $t^* = t$, and Eq. (31) gives $\lambda = 1$, which is *not* possible in case of pure shear deformation.

The coefficients R, R', R^* and $R^{*'}$ are determined by using the boundary conditions given by Eq. (24) and Eq. (26), with the second equation in Eq. (26) taking the form

$$(\lambda^4 + 2)\psi_{,112} - (2\lambda^2 + 1)\psi^*_{,112} + \psi_{,222} - \psi^*_{,222} - \overline{\rho}(\ddot{\psi}_{,2} - \ddot{\psi}^*_{,2}) = 0 \tag{32}$$

in this case, where $\overline{\rho}$ is given by Eq. (18). Substitution of ψ and ψ^* from Eq. (27)(with $m = 1$) and Eq. (28) in Eq. (24), the first equation in Eq. (26), and Eq. (32) leads to

$$\begin{aligned}
&1 + R + R' = R^* + R^{*'},\\
&t(1 - R) - iR' = t^* R^* + im^* R^{*'},\\
&(1 + R)(t^2 - 1) - 2R' = (t^{*2} - 1)R^* - (1 + m^{*2})R^{*'},\\
&t^{*2}\{2it(R - 1) + R'(t^2 - 1)\} + R^*(t^2 + t^{*2})it^* +\\
&R^{*'}\{t^{*4} + t^{*2}(m^{*2} - 1) - t^2\}m^* = 0.
\end{aligned} \tag{33}$$

In the latter equation use has been made of Eq. (31) in order to simplify the coefficients.

The solution of Eq. (33) may be written in the form

$$\begin{aligned}
R &= \frac{(t + i)(t^* - t)F(t)}{(t - i)(t^* + t)F(-t)},\\
R' &= \frac{2t(t - t^*)G'}{(t^* + i)(t - i)F(-t)},\\
R^* &= \frac{2t(t + i)G^*}{(i + t^*)(t^* - im^*)(t^* + t)F(-t)},\\
R^{*'} &= \frac{2t(i + t)(t^* - t)t^*}{i(m^* + it^*)F(-t)},
\end{aligned} \tag{34}$$

where $F(t)$, G', G^* are defined by

$$\begin{aligned}
F(t) &= t^2(t^* - i) - tt^* i(m^* + 1) + m^* t^{*2}(t^* + im^*),\\
G' &= m^* t^*(t^{*2} + im^* t^* + 1) - it^2,\\
G^* &= m^* t^{*4} + t^{*2}(t^2 + m^{*3} + m^{*2} + m^*) - m^* t^2,
\end{aligned}$$

and $F(-t)$ is obtained from $F(t)$ by replacing t by $-t$ without changing t^*. In these equations, for given t, m^* is obtained from Eq. (30) so as to have positive real part and t^* from Eq. (31). In Section 5 graphical results for the absolute values of R, $R^{'}$, R^* and $R^{*'}$ are given for illustration. All the figures have been produced using Mathematica [7].

5. Numerical Results

The slowness curves (superimposed) in Fig. 1(c) and Fig. 2(c) show that there is one reflected wave, one transmitted wave and two interfacial waves for each possible angle of incidence when $2\beta = \alpha + \gamma\,(x_2 < 0)$, $2\beta \neq \alpha + \gamma\,(x_2 > 0)$.

In Figs. 3-6, $|R|$, $|R'|$, $|R^*|$, $|R^{*'}|$ respectively are plotted, using Eq. (34), as functions of θ for a series of values of λ. Figs. (3-4) show, in particular, that as λ increases the maximum values of the reflected wave and the interfacial wave amplitudes (in $x_2 < 0$) $|R|$ and $|R'|$ increase.

In $x_2 > 0$, the character of the interfacial wave and the transmitted wave amplitudes $|R^*|$ and $|R^{*'}|$ is different (against different stretches).

For the grazing incidence ($\theta = 0$), from Eq. (31) we have

$$t^{*2} = \lambda^2(\lambda^2 - 2), \tag{35}$$

which shows that t^{*2} is positive when $\lambda > \sqrt{2}$ and negative for $\lambda < \sqrt{2}$. In Figs. 5(a-b) and Figs. 6(a-b) notice that $|R^*| \neq 0$ but $|R^{*'}| = 0$ when $\lambda < \sqrt{2}$ contrary to the results in Fig. 5(c-d) and Figs. 6(c-d) when $\lambda > \sqrt{2}$. In general the change in the amplitudes $|R^*|$ and $|R^{*'}|$ for $\lambda < \sqrt{2}$ and $\lambda > \sqrt{2}$ must be noted.

The graphical results show the general character of the effect of pure shear on the reflection and transmission of plane waves at the boundary of two half-spaces (corresponding to different strain-energy functions).

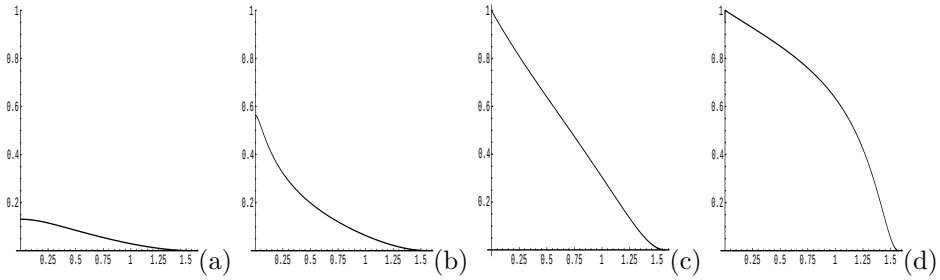


Figure 3. Plots of $|R|$ (reflected wave amplitude in $x_2 < 0$) against θ ($0 \leq \theta \leq \pi/2$) with the following values of λ: (a) 0.6, (b) 1.4, (c) 1.9, (d) 2.8.

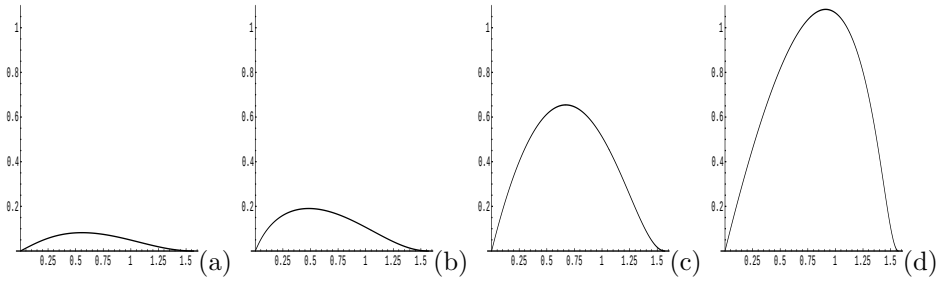


Figure 4. Plots of $|R'|$ (Interfacial wave amplitude in $x_2 < 0$) against θ ($0 \leq \theta \leq \pi/2$) with the following values of λ: (a) 0.6, (b) 1.4, (c) 1.9, (d) 2.8.

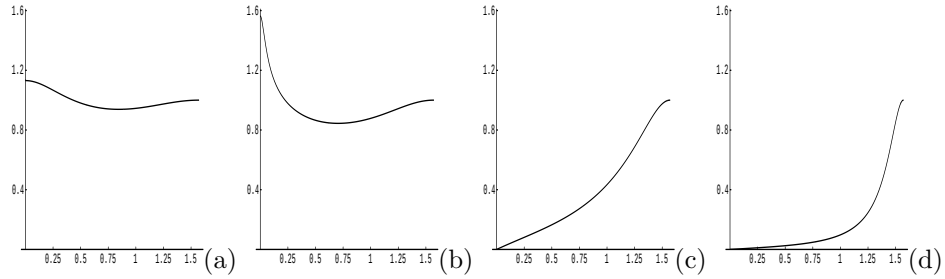


FIGURE 5. Plots of $|R^*|$ (Transmitted wave amplitude in $x_2 > 0$) against θ ($0 \leq \theta \leq \pi/2$) with the following values of λ: (a) 0.6, (b) 1.4, (c) 1.9, (d) 2.8.

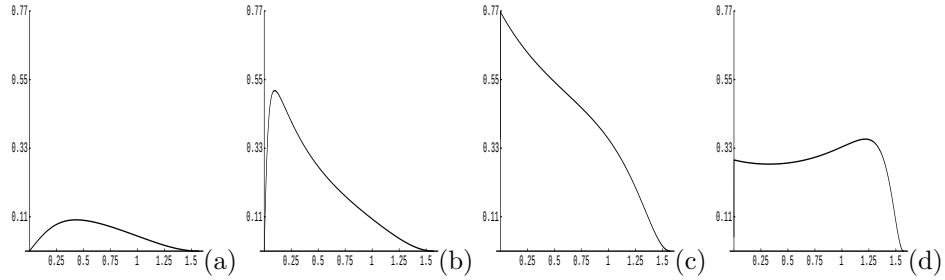


FIGURE 6. Plots of $|R^{*'}|$ (Interfacial wave amplitude in $x_2 > 0$) against θ ($0 \leq \theta \leq \pi/2$) with the following values of λ: (a) 0.6, (b) 1.4, (c) 1.9, (d) 2.8.

6. Conclusions

Since the angle of incidence θ is real, using first two equations in Eq. (34), the amplitudes of the reflected and interfacial waves are increasing smoothly (in $x_2 < 0$) by increasing the stretch as shown in Figs. (3-4).

As described in Section 5, at $\theta = 0$, $t^* = \tan\theta^*$ is real for $\lambda > \sqrt{2}$ from Eq. (35). Keeping this in view, the incident wave is not transmitted, as illustrated in Figs. 5(c-d). Do notice that interfacial wave is generated (in $x_2 > 0$) as shown in Figs. 6(c-d) for $\lambda > \sqrt{2}$.

Similarly θ^* is an imaginary angle from Eq. (35) for $\lambda < \sqrt{2}$. Therefore the behavior of $|R^*|$ ($|R^{*'}|$) at $\theta = 0$ for $\lambda < \sqrt{2}$ in Figs. 5(a-b) and Figs. 6(a-b) respectively, is similar to that of $|R^{*'}|$ ($|R^*|$) at $\theta = 0$ for $\lambda > \sqrt{2}$, as graphically shown in Figs. 6(c-d) and Figs. 5(c-d) respectively.

By comparing with the experimental data, the results described here, provide theoretical basis for characterization of material properties of rubber like solids i.e. stretches, strain-energy functions and elastic moduli. One can picture situations of practical interest where a solid is stretched and sheared: for instance, a rubber isolator under a bridge is subjected to vertical compression and then is sheared as a result of thermal extensions and contractions of the roadway.

The method and the results presented in the paper can be *extended* to the wave propagation problem by interchanging the strain-energy functions of two half-spaces i.e.

$$2\beta \neq \alpha + \gamma \, (x_2 < 0), \text{ and } 2\beta = \alpha + \gamma \, (x_2 > 0).$$

With reference to Fig. 1 problem would be similar i.e. there will be one reflected wave, one transmitted wave along with two interfacial waves *but* according to the Fig. 2, (in addition to Fig. 1 option) there might be *two* reflected waves (in $x_2 < 0$) with two interfacial waves (in $x_2 > 0$) with no transmitted wave for a certain range of the angle of incidence. Propagation of two interfacial waves (in the upper half-space), to best of the knowledge of author, *never* appeared in *linear* elasticity!

Another possible extension is to study the reflection and transmission of waves in unconstrained (*compressible)* elastic solids. See, for example [3], where reflection of plane waves from the boundary of a pre-stressed compressible elastic half-space is studied.

References

[1] W. Hussain and R. W. Ogden, On the reflection of plane waves at the boundary of an elastic half-space subject to simple shear. *Int. J. Engng Sci.* **37**, 1999, 1549-1576.

[2] R. W. Ogden and D. A. Sotiropoulos, The effect of pre-stress on the propagation and reflection of plane waves in incompressible elastic solids. *IMA J. Appl. Math.* **59**, 1997, 95-121.

[3] R. W. Ogden and D. A. Sotiropoulos, Reflection of plane waves from the boundary of a pre-stressed compressible elastic half-space. *IMA J. Appl. Math.* **61**, 1998, 61-90.

[4] S. Dey and S. K. Addy, Reflection and refraction of plane waves under initial stresses at an interface. *Int. J. Nonlinear Mech.* **14**, 1979, 101-110.

[5] A. N. Norris, Propagation of plane waves in a pre-stressed elastic medium. *J. Acoust. Soc. Am.* **74**, 1983, 1642-1643.

[6] R. W. Ogden, Non-linear Elastic Deformations, Dover Publications, Inc. Mineola, New York, 1997.

[7] S. Wolfrom, Mathematica, version 5, Wolfrom Research, Champaign, Illinois, 2003.

Department of Mathematics, School of Arts and Sciences, Lahore University of Management Sciences (LUMS), Opposite Sector 'U', D.H.A., Lahore Cantt. 54792, Pakistan.

E-mail address: wasiq@lums.edu.pk

ALBANIAN JOURNAL
OF MATHEMATICS
Volume 1, Number 2, Pages 99–114
ISSN 1930-1235: (electronic version)

EXACT ASYMPTOTICS FOR TYPE I BIVARIATE ELLIPTICAL DISTRIBUTIONS

ENKELEJD HASHORVA

ABSTRACT. Let (S_1, S_2) be a bivariate spherical random vector with associated random radius which has distribution function in the Gumbel max-domain of attraction. In this paper we obtain an exact asymptotic expansion of the tail probability $\boldsymbol{P}\{S_1 > u_n, \rho_n S_1 + \sqrt{1-\rho_n^2} S_2 > v_n\}, \rho_n \in (-1,1)$ with $u_n, v_n, n \geq 1$ constants letting $u_n \to \infty$ and $\rho_n \to \rho \in (-1,1)$. As an application of our result the limit distribution of the joint and the partial excess distribution is obtained.

Dedicated to Professor Jürg Hüsler
on the Occasion of his 60th Birthday

1. INTRODUCTION

Let (S_1, S_2) be a spherical random vector with associated random radius $R := \sqrt{S_1^2 + S_2^2} > 0$ almost surely. Basic properties of spherical random vectors are obtained in Cambanis et al. (1981). So if $R > 0$ almost surely, then we have the stochastic representation

$$(S_1, S_2) \stackrel{d}{=} (RO_1, RO_2),$$

with (O_1, O_2) uniformly distributed on the unit circle of $\mathbb{R}^2$ being further independent of R ($\stackrel{d}{=}$ stands for equality of distribution functions).

A natural generalisation of this class is the class of elliptical random vectors, defined as linear combination of spherical random vectors. Elliptical random vectors are both from the theoretical and the practical point of view very interesting. This class is very large, including the prominent Gaussian and Kotz distribution.

Throughout this paper we consider elliptical random vectors $(X_0, Y_0), (X_1, Y_1), \ldots$ in $\mathbb{R}^2$ with stochastic representation

$$(X_n, Y_n) \stackrel{d}{=} (S_1, \rho_n S_1 + \sqrt{1-\rho_n^2} S_2), \quad \rho_n \in (-1,1), \quad n \geq 0. \tag{1}$$

The basic distribution properties of elliptical random vectors are well-known, see e.g., Kotz (1975), Cambanis et al. (1981), Anderson and Fang (1990), Fang et. al (1990), Fang and Zhang (1990), Szabłowski (1990), Berman (1992), Gupta and Varga (1993), Kano (1994), Kotz and Ostrovskii (1994) among several others.

The main asymptotic properties of bivariate elliptical random vectors are derived

Received by the editors February 11, 2007 and, in revised form, May 21, 2007.
2000 *Mathematics Subject Classification.* Primary 60F05; Secondary 60G70.
Key words and phrases. Elliptical random vectors, exact asymptotics, Gumbel max-domain of attraction, conditional limiting theorems, excess distribution.

by Berman (1982,1983) culminating in his excellent monograph Berman (1992). Berman's focus was the asymptotic properties of the Berman processes. The work of Berman has been therefore not referred for a long time in the literature of multivariate distributions. Similar results for bivariate spherical random vectors are obtained in Carnal (1970), Gale (1980), Eddy and Gale (1981) in the context of convex hull asymptotics.

In this paper we are interested in the exact asymptotics of the tail probability

$$\boldsymbol{P}\{X_n > u_n, Y_n > v_n\} = \quad u_n, v_n \in \mathbb{R}, n \geq 1 \tag{2}$$

letting u_n tend to ∞.

Intuitively, since the associated random radius R is the only unknown component of the elliptical random vectors, we expect that its tail asymptotic behaviour determines the asymptotic behaviour of (2). This is the case for the Gaussian random vectors (see e.g., Hashorva and Hüsler (2003) or Hashorva (2005a)).

Indeed the Gaussian case has been treated in very many papers. The result for the case $u_n = v_n, n \geq 1$ is given in Berman (1962). See Dai and Mukherjea (2001) or Hashorva (2005a) for further references.

The square of the associated random radius of a d-dimensional Gaussian vector is chi-squared distributed with d degrees of freedom. From the extreme value theory we know that R in the Gaussian case has distribution function F in the max-domain of attraction of the Gumbel distribution function $\Lambda(x) := \exp(-\exp(-x)), x \in \mathbb{R}$. Motivated by this fact, in the recent paper Hashorva (2006b) an asymptotic expansion of the tail probability for a general multivariate setup is obtained. Those results can be applied to our case when ρ_n does not depend on n.

Making use of a tractable formula for the bivariate elliptical distributions we obtain in this paper the asymptotic expansion of the tail probability of interest allowing ρ_n to depend on n, and provide a simpler proof than that in the aforementioned paper.

Further, we apply our result to study the asymptotics of bivariate excess distributions.

2. Preliminaries

In this section we present some standard notation and give few preliminary results. The main results are given in Section 3, followed by the proofs in Section 4 (last one).

Given a random variable Y with distribution function H, we shall denote this alternatively as $Y \sim H$. If F is the Gamma distribution with positive parameters a, b we write $Y \sim Gamma(a, b)$.

Next, let $(X_n, Y_n), n \geq 0$ be a bivariate elliptical random vector as in (1), and write throughout this paper $(X, Y), \rho$ instead of $(X_0, Y_0), \rho_0$.

We assume in the following that the associated random radius R has distribution function F such that $F(0) = 0$. Further, we impose a certain asymptotic restriction on the distribution function F, namely we suppose that there exists a positive scaling function w such that

$$\lim_{u \uparrow x_F} \frac{1 - F(u + x/w(u))}{1 - F(u)} \quad = \quad \exp(-x), \quad \forall x \in \mathbb{R} \tag{3}$$

is valid with $x_F \in (0, \infty]$ the upper endpoint of F. The above condition is equivalent (see the standard monographs de Haan (1970), Leadbetter et al. (1983), Galambos

(1987), Resnick (1987), Reiss (1989), Falk et al. (2004) or Kotz and Nadarajah (2005), or de Haan and Ferreira (2006)) with the fact that F is in the Gumbel max-domain of attraction, meaning the sample maxima of a random sample with underlying distribution function F converges in distribution (after an affine normalisation) to a Gumbel random variable.
We refer to (X, Y) in the case F satisfies (3) as Type I elliptical random vector.
The scaling function w can be defined by

$$w(u) \quad := \quad \frac{1-F(u)}{\int_u^{x_F}[1-F(s)]\,ds}, \quad u \in (0, x_F]. \tag{4}$$

Further, uniformly on the compact sets of $z \in \mathbb{R}$

$$\lim_{u \uparrow x_F} \frac{w(u + z/w(u))}{w(u)} \quad = \quad 1, \tag{5}$$

and

$$\lim_{u \uparrow x_F} uw(u) \quad = \quad \infty. \tag{6}$$

In view of Lemma 6.2 of Berman (1982) (given also in Lemma 12.1.2 in Berman (1992))

$$aS_1 + bS_2 \stackrel{d}{=} \sqrt{a^2+b^2}S_1, \quad \forall a, b \in \mathbb{R},$$

hence for $(X_n, Y_n) \stackrel{d}{=} (S_1, \rho_n S_1 + \sqrt{1-\rho_n^2}S_2)$ with $\rho_n \in (-1, 1), n \geq 0$ (as in (1)) we have

$$X_n \stackrel{d}{=} Y_n \stackrel{d}{=} S_1. \tag{7}$$

Applying Theorem 12.3.1 of Berman (1992) we obtain $(n \to \infty)$

$$\boldsymbol{P}\{X > u_n\} = \boldsymbol{P}\{S_1 > u_n\} \quad = \quad (1+o(1))\left(\frac{1}{u_n w(u_n)}\right)^{1/2} \frac{1}{\sqrt{2\pi}}[1-F(u_n)], \tag{8}$$

provided that $\lim_{n\to\infty} u_n = x_F \in (0, \infty]$ and F satisfies (3) with the scaling function w. Consequently, when (X_n, Y_n) is a Type I elliptical random vector, then the asymptotic tail behaviour of its components is known. In the special case that (X_n, Y_n) has independent components we have

$$\boldsymbol{P}\{X_n > u_n, Y_n > v_n\} = \boldsymbol{P}\{X_n > u_n\}\boldsymbol{P}\{Y_n > v_n\}, \quad n \geq 1,$$

hence for this instance there is nothing to investigate.
Provided that (X_n, Y_n) has a density function, we know that X_n and Y_n are independent (see e.g., Fang et al. (1990), Hashorva et al. (2007)) only when X_n and Y_n are standard Gaussian random variables. Therefore the above simplification of our problem of interest is only possible for a trivial case. In the case (X_n, Y_n) is Gaussian and X_n, Y_n are correlated ($\rho_n \neq 0$), the exact asymptotics of the probability of interest is know (see e.g., Hashorva (2005a)). The general elliptical case is derived in Hashorva (2006b).
Next, we consider briefly the bivariate Gaussian case and then give a conditional limiting result which will be utilised in Section 3. Assume for simplicity that

$$v_n = au_n, \quad a \in (-\infty, 1], \quad n \geq 1,$$

and $u_n, n \geq 1$ is a positive sequence converging to infinity.

It turns out that the correlation ρ plays via the Savage condition (see e.g., Hashorva and Hüsler (2003)) a crucial role in determining the joint tail asymptotic behaviour of (X,Y). In the bivariate case this condition is very simple to formulate, namely if $a>\rho$ we have

$$\boldsymbol{P}\{X>u_n, Y>au_n\} \quad = \quad (1+o(1))C_{a,\rho}\frac{\exp(-(u_n\boldsymbol{\alpha}_{a,\rho})^2/2)}{2\pi u_n^2}, \quad n\to\infty,$$

with

$$\boldsymbol{\alpha}_{a,\rho}:=\sqrt{(1-2a\rho+a^2)/(1-\rho^2)}>1,\ C_{a,\rho} \quad := \quad \frac{(1-\rho^2)^{3/2}}{(1-a\rho)(a-\rho)}>0. \tag{9}$$

If $a\le\rho$ then

$$\boldsymbol{P}\{X>u_n, Y>au_n\} \quad = \quad (1+o(1))\mathbf{1}_{\rho,a}\frac{\exp(-u_n^2/2)}{\sqrt{2\pi}u_n}, \quad n\to\infty \tag{10}$$

is valid with $\mathbf{1}_{\rho,a}:=1/2$ if $\rho=a$ and $\mathbf{1}_{\rho,a}:=1$, otherwise.
If the Savage condition holds, i.e., $\rho>a$ then $\boldsymbol{\alpha}_{a,\rho}>1$, implying that the joint tail asymptotics is faster than the convergence rate to 0 of $\boldsymbol{P}\{X>u_n\}$. Moreover, the speed of the convergence is governed by $\boldsymbol{\alpha}_{a,\rho}$, which is actually the attained minimum of a related quadratic programming problem (see e.g., Hashorva (2005a)).
If the Savage condition does not hold, then (10) shows that the asymptotics is of the same rate as of $\boldsymbol{P}\{X>u_n\}, n\to\infty$. The later asymptotics is well-known and related to Mills Ratio (see e.g., Berman (1962)).
In the Gaussian case (3) holds with $w(t)=t, t>0$, hence we may write (10) using further (8)

$$\begin{aligned}&\boldsymbol{P}\{X>u_n, Y>au_n\}\\ &\quad = \quad (1+o(1))\mathbf{1}_{\rho,a}\boldsymbol{P}\{X>u_n\}\\ &\quad = \quad (1+o(1))\mathbf{1}_{\rho,a}\left(\frac{1}{u_n w(u_n)}\right)^{1/2}\frac{1}{\sqrt{2\pi}}[1-F(u_n)], \quad n\to\infty\end{aligned}$$

showing that the asymptotics is defined by $1-F(u_n)$ and $u_n w(u_n)$. This is the case for Type I elliptical random vectors in general as shown in Hashorva (2006b) (corresponding to our case ρ not depending on n). We shall present in this paper another proof of that result (see Theorem 2 below), and consider further the case ρ_n depends on n.
Finally for ease of reference we present next a conditional limiting theorem proved in Theorem 4.1 of Berman (1983) (see also Theorem 12.4.1 of Berman (1992)). That result first appears in Lemma 8.2 of Berman (1982) (with some additional restrictions). More special case are dealt with in Gale (1980), Eddy and Gale (1981). See for details Abdous et al. (2005), Abdous et al. (2006), Hashorva (2006a), or Hashorva et al. (2007).
Recent deep articles on the subjects are Heffernan and Tawn (2004), Butler and Tawn (2005) and Heffernan and Resnick (2005).

We denote throughout the paper a Gaussian random variable with mean 0 and variance $\sqrt{1-\rho^2}\in(0,1]$ by Z_ρ, i.e.,

$$Z_\rho \quad \stackrel{d}{=} \quad \sqrt{1-\rho^2}W, \tag{11}$$

where W is a standard Gaussian random variable.

Theorem 1. *[Berman (1992)] Let (S_1, S_2) be a spherical bivariate random vector with associated random radius $R := \sqrt{S_1^2 + S_2^2} > 0$ almost surely. If the distribution function F of R satisfies* (3) *with the scaling function w, then we have for any $\rho \in (-1, 1)$ and $u_n < x_F, n \geq 1$ such that $\lim_{n\to\infty} u_n = x_F$*

$$q_n\Big(\rho S_1 + \sqrt{1-\rho^2}S_2 - \rho u_n\Big)\big|S_1 > u_n \quad \stackrel{d}{\to} \quad Z_\rho, \quad n \to \infty, \tag{12}$$

with $q_n := \sqrt{w(u_n)/u_n}$ and Z_ρ as in (11).

For the case $\rho = 0$ the proof is given in Theorem 12.4.1 of Berman (1992). The case $\rho \in (-1, 1)$ is proved in Berman (1992) in Theorem 12.5.1 (see (12.5.5)). In fact from (12.5.7) therein we have the convergence in probability

$$q_n|S_1 - u_n|\big|S_1 > u_n \quad \stackrel{p}{\to} \quad 0, \quad n \to \infty, \tag{13}$$

hence the proof of the case $\rho \neq 0$ is a simple consequence of Theorem 12.4.1 of Berman (1992) and (13).

3. Main Results

In this section we consider bivariate elliptical random vectors $(X, Y), (X_1, X_2), \ldots$ with stochastic representation (1) and associated random radius $R \sim F$. We consider for simplicity only the case F has an infinite upper endpoint.
Given two sequences $u_n, v_n, n \geq 1$ we derive in the main result below an asymptotic expansion for $\boldsymbol{P}\{X_n > u_n, Y_n > v_n\}$ letting u_n tend to ∞. For $v_n, n \geq 1$ we require that $\lim_{n\to\infty} v_n/u_n = a \in (-\infty, 1]$. As illustrated by the Gaussian example above the pseudo-correlation coefficient ρ (recall (1)) plays a central role for the asymptotics via the Savage condition.
The main assumption in this section is that (X_n, Y_n) is a Type I elliptical random vector, i.e., the distribution function F of the associated random radius R is in the Gumbel max-domain of attraction.

Theorem 2. *Let $(X, Y), (X_1, Y_1), \ldots$ be Type I bivariate elliptical random vector with stochastic representation* (1)*, where $\rho, \rho_n \in (-1, 1), n \geq 1$, and let $u_n, v_n \in \mathbb{R}, n \geq 1$ be given constants such that $\lim_{n\to\infty} u_n = \infty$. Assume that the associated random radius $R \sim F$ is almost surely positive with F in the Gumbel max-domain of attraction satisfying* (3) *with the positive scaling function w, and upper endpoint $x_F = \infty$. Suppose further that $\lim_{n\to\infty} \rho_n = \rho \in (-1, 1)$ and let Z_ρ be as in* (11).
i) If for some $z \in [-\infty, \infty)$

$$\lim_{n\to\infty} q_n[v_n - \rho_n u_n] = z \tag{14}$$

holds with $q_n := \sqrt{w(u_n)/u_n}, n \geq 1$, then for any sequence $y_n \in \mathbb{R}, n \geq 1$ such that $\lim_{n\to\infty} y_n = y \in [-\infty, \infty)$

$$\begin{aligned} &\boldsymbol{P}\{X_n > u_n, Y_n > v_n + y_n/q_n\} \\ &= (1 + o(1))\boldsymbol{P}\{Z_\rho > y + z\}\frac{1}{\sqrt{2\pi}}\left(\frac{1}{u_n w(u_n)}\right)^{1/2}[1 - F(u_n)] \qquad (15) \\ &= (1 + o(1))\boldsymbol{P}\{Z_\rho > y + z\}\boldsymbol{P}\{X > u_n\} \qquad (16) \end{aligned}$$

holds as $n \to \infty$.
ii) Set $a_n := v_n/u_n, n \geq 1$ *and suppose further that* $a_n \in (\rho_n, 1]$ *for all large* n *and*

$$\lim_{n\to\infty} a_n \quad = \quad a \in (\rho, 1]. \tag{17}$$

Then we have

$$\begin{aligned} & \boldsymbol{P}\{X_n > u_n, Y_n > v_n\} \\ &= (1+o(1))\frac{\boldsymbol{\alpha}_{a,\rho} C_{a,\rho}}{2\pi} \frac{1}{u_n w(u_n^*)}[1-F(u_n^*)] \qquad (18) \\ &= (1+o(1))\frac{\boldsymbol{\alpha}_{a,\rho}^2 C_{a,\rho}}{\sqrt{2\pi}} \left(\frac{1}{u_n^* w(u_n^*)}\right)^{1/2} \boldsymbol{P}\{X > u_n^*\}, \quad n \to \infty, \qquad (19) \end{aligned}$$

with $\boldsymbol{\alpha}_{a,\rho}, C_{a,\rho}$ *as in* (9) *and* $u_n^* := \boldsymbol{\alpha}_{n,a,\rho} u_n, n \geq 1$ *where*

$$\boldsymbol{\alpha}_{n,a,\rho} := \sqrt{(1-2a_n\rho_n + a_n^2)/(1-\rho^2)} \quad \to \quad \boldsymbol{\alpha}_{a,\rho} > 1, \quad n \to \infty. \tag{20}$$

Remarks 1. a) Since F is in the Gumbel max-domain of attraction we have

$$\lim_{t\to\infty} \frac{1-F(ct)}{1-F(t)} = 0, \quad \forall c > 1,$$

hence (8) yields also

$$\lim_{t\to\infty} \frac{\boldsymbol{P}\{X_n > ct\}}{\boldsymbol{P}\{X_n > t\}} = \lim_{t\to\infty} \frac{\boldsymbol{P}\{S_1 > ct\}}{\boldsymbol{P}\{S_1 > t\}} = 0, \quad \forall c > 1.$$

Consequently by (20)

$$\lim_{n\to\infty} \frac{\boldsymbol{P}\{X_n > u_n^*\}}{\boldsymbol{P}\{X_n > u_n\}} = 0.$$

Further (6) implies

$$\lim_{n\to\infty} u_n w(u_n) = \lim_{n\to\infty} u_n w(u_n^*) \quad = \quad \infty, \tag{21}$$

hence the asymptotics in (18) is faster than the one in (15).
b) If the distribution function F has a finite upper endpoint $x_F \in (0, \infty)$, then the first statement above still holds for $u_n \to x_F$ as $n \to \infty$ and $v_n, n \geq 1$ satisfying further

$$u_n^2 - 2\rho_n u_n v_n + v_n^2 \quad < \quad 1 - \rho_n^2, \quad n \geq 1. \tag{22}$$

c) Our asymptotic results in the above theorem confirm (for the case $\rho_n = \rho, n \geq 1$) the ones previously obtained in Hashorva (2006b).
Note that if ρ_n depends on n, then the rate of convergence in (18) depends explicitly on ρ_n. Furthermore, the conditions leading to both statements above need to be formulated with ρ_n instead of ρ (see (14)).

In the special case (which is common in applications), namely $v_n = u_n a, \rho_n = \rho, n \geq 1$ with $a \in (-\infty, 1]$ and $\rho \in (-1, 1)$ we have:

Corollary 3. *Under the assumptions and the notation of Theorem 2, if further* $a_n = a \in (-\infty, 1]$ *for all large* n *then we have:*
i) In the case $a < \rho$

$$\lim_{n\to\infty} \frac{\boldsymbol{P}\{X > u_n, Y > au_n + y_n/q_n\}}{\boldsymbol{P}\{X > u_n\}} \quad = \quad 1. \tag{23}$$

ii) In the case $a = \rho$

$$\lim_{n\to\infty} \frac{\boldsymbol{P}\{X > u_n, Y > au_n + y_n/q_n\}}{\boldsymbol{P}\{X > u_n\}} \quad = \quad \boldsymbol{P}\{Z_\rho > y\}. \tag{24}$$

iii) If $a \in (\rho, 1]$ *then we have as* $n \to \infty$

$$\begin{aligned} &\boldsymbol{P}\{X > u_n, Y > au_n\} \\ &= (1 + o(1))\frac{\boldsymbol{\alpha}_{a,\rho}^{3/2} C_{a,\rho}}{\sqrt{2\pi}} \left(\frac{1}{u_n w(\boldsymbol{\alpha}_{a,\rho} u_n)}\right)^{1/2} \boldsymbol{P}\{X > \boldsymbol{\alpha}_{a,\rho} u_n\}, \end{aligned} \tag{25}$$

with $C_{a,\rho}, \boldsymbol{\alpha}_{a,\rho}$ *as in Theorem 2.*

Corollary 4. *Under the assumptions of Theorem 2 we have:*
i) In the case $\rho \in (0, 1)$

$$\boldsymbol{P}\{X > u_n, Y > y\} \quad = \quad (1 + o(1))\boldsymbol{P}\{X > u_n\}, \quad n \to \infty \tag{26}$$

is valid for any $y \in \mathbb{R}$.
ii) In the case $\rho = 0$ *and*

$$\lim_{n\to\infty} \left(\frac{w(u_n)}{u_n}\right)^{1/2} \quad = \quad b \in [0, \infty), \tag{27}$$

we have for $y \in \mathbb{R}$

$$\boldsymbol{P}\{X > u_n, Y > y\} \quad = \quad (1 + o(1))\boldsymbol{P}\{Z_\rho > by\}\boldsymbol{P}\{X > u_n\}, \quad n \to \infty. \tag{28}$$

iii) If $\rho \in (-1, 0)$ *then for any* $y > 0$

$$\begin{aligned} &\boldsymbol{P}\{X > u_n, Y > y\} \\ &= (1 + o(1))\left(\frac{(1-\rho^2)^3}{2\pi\rho^2}\right)^{1/2} \left(\frac{1}{u_n w(u_n^*))}\right)^{1/2} \boldsymbol{P}\{X > u_n^*\}, \quad n \to \infty, \end{aligned} \tag{29}$$

is valid with

$$u_n^* := \sqrt{(y^2 - 2\rho u_n y + u_n^2)/(1-\rho^2)}, \quad n \in \mathbb{N}.$$

Remark 1. The above corollary is important also in a distributional context. In view of (26) and (29), X and Y cannot be independent if (X, Y) is a Type I elliptical random vector with pseudo-correlation $\rho \in (-1, 1), \rho \neq 0$.
When X, Y are independent with (X, Y) Type I, then we have thus $\rho = 0$, hence if further (27) holds, then (28) implies that $Y \stackrel{d}{=} Z_\rho/b$.

We consider next an illustrating example:
Example 1. [**Kotz Type III**] Let $(X, Y) = R(O_1, \rho O_1 + \sqrt{1-\rho^2}O_2)$, with R a positive random radius independent of the bivariate random vector (O_1, O_2) which is uniformly distributed on the unit circle of $\mathbb{R}^2$. We call $\boldsymbol{X}$ a Kotz Type III elliptical random vector if further

$$\boldsymbol{P}\{R > u\} \quad = \quad (1 + o(1))Ku^N \exp(-ru^\delta), \quad K > 0, \delta \in \mathbb{R}, N \in \mathbb{R}, \quad u \to \infty,$$

with $\delta \leq 0$ if $N < 0$. We consider next only the case $\delta > 0$. Define the function w by

$$w(u) \quad = \quad r\delta u^{\delta - 1}, \quad u > 0.$$

For any $x \in \mathbb{R}$ we have

$$\frac{\boldsymbol{P}\{R > u + x/w(u)\}}{\boldsymbol{P}\{R > u\}} = (1 + o(1)) \exp\Big(-ru^{\delta}\Big[\Big(1 + \frac{x}{r\delta u^{\delta}}\Big)^{\delta} - 1\Big]\Big) \to \exp(-x)$$

as $u \to \infty$, implying that F is in the Gumbel max-domain of attraction with the scaling function w. In view of (8) we have

$$\boldsymbol{P}\{X > u\} \quad = \quad (1 + o(1)) \frac{K}{\sqrt{2r\delta\pi}} u^{N-\delta/2} \exp(-ru^{\delta}), \quad u \to \infty.$$

Let $u_n \to \infty$ and $y_n \to y \in \mathbb{R}$ be two given sequence and let $a \in (-\infty, 1]$ be a given constant. Then by the above results we have if $\rho < a$

$$\begin{aligned} &\boldsymbol{P}\Big\{X > u_n, Y > au_n + y_n\sqrt{r\delta u_n^{\delta-1}/u_n}\Big\} \\ &= \quad (1 + o(1)) \frac{K}{\sqrt{2r\delta\pi}} u_n^{N-\delta/2} \exp(-ru_n^{\delta}), \quad n \to \infty. \end{aligned}$$

If $a = \rho$ similar asymptotics follows where the constant is additionally multiplied by $\boldsymbol{P}\{Z_\rho > y\}$.
Assuming that $a \in (\rho, 1]$ we obtain as $n \to \infty$

$$\boldsymbol{P}\{X > u_n, Y > au_n\} \quad = \quad (1 + o(1)) \frac{K\boldsymbol{\alpha}_{a,\rho}^{2-\delta+N} C_{a,\rho}}{2r\delta\pi} u_n^{N-\delta} \exp(-r(\boldsymbol{\alpha}_{a,\rho} u_n)^{\delta}).$$

Note in passing that the Gaussian case corresponds to the choice of parameters

$$K = 1, N = 0, \delta = 2, r = 1/2.$$

4. Approximation of Excess Distribution

Consider $(X, Y), (X_1, Y_1), \ldots$ Type I elliptical bivariate random vector with stochastic representation (1) and associated random radius $R \sim F$. Let $u_n, n \geq 1$ be a positive sequence such that $\lim_{n\to\infty} u_n = x_F, |u_n| < x_F, n \geq 1$. The random variable $X_n - u_n | X_n > u_n$ is the excess of X_n above the threshold u_n given X_n jumps the threshold.
An immediate consequence of the assumption F is in the Gumbel max-domain of attraction with the positive scaling function w is the convergence in distribution of the corresponding excess random variables above the threshold u_n

$$w(u_n)(X_n - u_n)|X_n > u_n \xrightarrow{d} U, \quad w(u_n)(Y_n - u_n)|Y_n > u_n \xrightarrow{d} U, \tag{30}$$

with $U \sim Gamma(1, 1)$ a unit Exponential random variable.

Another interesting situation arises when we additionally condition on the other component being large, i.e., considering the joint excess bivariate random sequence (with respect to $u_n, v_n, n \geq 1$)

$$(X_{u_n,Y,v_n}, Y_{v_n,X,u_n}) \quad := \quad (X_n - u_n, Y_n - v_n)\big|X_n > u_n, Y_n > v_n, \quad n \geq 1.$$

With the above notation we can re-write (12) and (13) as

$$\Big(q_n(Y_n - \rho_n u_n), q_n(X_n - u_n)\Big)\Big|X_n > u_n \quad \xrightarrow{d} \quad (Z_\rho, 0\Big), \quad n \to \infty, \tag{31}$$

where $q_n := \sqrt{w(u_n)/u_n}, n \geq 1$ and Z_ρ as in (11).
Convergence in distribution is stated in the next theorem, which is a slight modification of Berman's result presented in the previous section.

Theorem 5. *Let (S_1, S_2) be a Type I bivariate spherical random vector with associated random radius $R \sim F$, where F satisfies (3) with upper endpoint $x_F \in (0, \infty]$ and the scaling function w. Let $u_n < x_F, n \geq 1, \rho_n \in (-1, 1)$ be constants such that $\lim_{n\to\infty} u_n = x_F$ and set*

$$X_n := S_1, \quad Y_n := \rho_n S_1 + \sqrt{1-\rho_n^2} S_2, n \geq 1.$$

If $\lim_{n\to\infty} \rho_n = \rho \in (-1, 1)$, then we have the convergence in distribution $(n \to \infty)$

$$\Big(q_n\big(Y_n - \rho_n u_n\big), w(u_n)(X_n - u_n)\Big)\Big| X_n > u_n \quad \stackrel{d}{\to} \quad \big(Z_\rho, U\big), \quad n \to \infty, \tag{32}$$

where $q_n := \sqrt{w(u_n)/u_n}, n \geq 1$, and Z_ρ as in (11) independent of $U \sim Gamma(1, 1)$.

Hashorva (2006b) obtains in Theorem 5.1 the convergence in distribution of the joint excess sequence $(X_{u_n,Y,u_n}, Y_{u_n,X,u_n}), n \geq 1$. A similar (independent) result appears in Asimit and Jones (2007) under the further restriction that the scaling function w is regularly varying and F has an infinite upper endpoint.
We apply our previous results to derive several approximations in the next theorem.

Theorem 6. *Let $F, (X, Y), \rho, Z_\rho, (X_n, Y_n), \rho_n, a_n, u_n, v_n, n \geq 1$ be as in Theorem 2, F satisfies (3) with the scaling function w and upper endpoint $x_F = \infty$, and let $h_{ni}, n \geq 1, i = 1, 2$ be positive constants such that*

$$\lim_{n\to\infty} q_n h_{ni} \quad = \quad c_i \in [0, \infty), \quad i = 1, 2, \tag{33}$$

with $q_n := \sqrt{w(u_n)/u_n}, n \geq 1$.
i) If (14) holds with $z \in [-\infty, \infty)$, then we have for any $x, y \in \mathbb{R}$

$$\lim_{n\to\infty} \frac{\boldsymbol{P}\{X_{u_n,Y,v_n} > h_{n1}x, Y_{v_n,X,u_n} > h_{n2}y\}}{\boldsymbol{P}\{X > u_n + h_{n1}x | X > u_n\}} \quad = \quad \overline{\Phi}_{\rho,z}(c_2 y - \rho c_1 x), \tag{34}$$

where $\overline{\Phi}_{\rho,z}(s) := \boldsymbol{P}\{Z_\rho > s + z\}/\boldsymbol{P}\{Z_\rho > z\}, s \in \mathbb{R}$.
Furthermore, in the case that $z \in \mathbb{R}$ we have the convergence in distribution

$$\Big(w(u_n) X_{u_n,Y,v_n}, q_n Y_{v_n,X,u_n}\Big) \quad \stackrel{d}{\to} \quad (U, V_z), \quad n \to \infty, \tag{35}$$

with $U \sim Gamma(1, 1), V_z \sim 1 - \overline{\Phi}_{\rho,z}$.
ii) Set $u_n^* := \sqrt{(u_n^2 - 2\rho_n u_n v_n + v_n^2)/(1-\rho_n^2)}, n \in \mathbb{N}$. If further (17) is satisfied, we then have the convergence in distribution

$$\Big(w(u_n^*) X_{u_n,Y,v_n}, w(u_n^*) Y_{v_n,X,u_n}\Big) \quad \stackrel{d}{\to} \quad (U_1, U_2), \quad n \to \infty, \tag{36}$$

where

$$U_1 \sim Gamma(1, \frac{1 - a\rho}{\boldsymbol{\alpha}_{a,\rho}(1-\rho^2)}), \quad U_2 \sim Gamma(1, \frac{a - \rho}{\boldsymbol{\alpha}_{a,\rho}(1-\rho^2)}),$$

with U_1, U_2 being further independent and $\boldsymbol{\alpha}_{a,\rho} := \sqrt{(1 - 2a\rho + a^2)/(1-\rho^2)} > 1$.

We give next an illustrating example.
Example 2. Let $X, Y, \rho, F, w, u_n, u_n^*, v_n, n \geq 1$ be as in Theorem 6. Assume that for all $c > 1$

$$\lim_{u\to\infty} \frac{w(cu)}{w(u)} \quad = \quad c^\lambda, \quad \lambda \in (-1, \infty). \tag{37}$$

If (17) holds, then we obtain (the convergence above is locally uniformly)

$$\lim_{n\to\infty}\frac{w(u_n^*)}{w(u_n)}=\frac{w(u_n\boldsymbol{\alpha}_{a,\rho})}{w(u_n)}=\boldsymbol{\alpha}_{a,\rho}^{\lambda}>1,$$

hence for any x,y positive

$$\begin{aligned}\lim_{n\to\infty}\boldsymbol{P}\Big\{X>u_n+\frac{x}{w(u_n^*)},Y>v_n+\frac{y}{w(u_n^*)}\big|X>u_n,Y>v_n\Big\}\\ =\lim_{n\to\infty}\boldsymbol{P}\Big\{X>u_n+\frac{\boldsymbol{\alpha}_{a,\rho}^{-\lambda}x}{w(u_n)},Y>v_n+\frac{\boldsymbol{\alpha}_{a,\rho}^{-\lambda}y}{w(u_n)}\big|X>u_n,Y>v_n\Big\}\\ =\exp\Big(-\boldsymbol{\alpha}_{a,\rho}^{-\lambda}([\overline{K}_{a,\rho}x+K_{a,\rho}y]\Big)\\ =\exp\Big(-\frac{1-a\rho}{\boldsymbol{\alpha}_{a,\rho}^{\lambda+1}(1-\rho^2)}x+\frac{a-\rho}{\boldsymbol{\alpha}_{a,\rho}^{\lambda+1}(1-\rho^2)}y\Big).\end{aligned}$$

Note in passing that if (X,Y) is a Kotz Type III elliptical random vector with $\delta>0$ then (37) holds with $\lambda=\delta-1$.

5. Proofs

For the proof of Theorem 2 we need the next lemma, which could be of some interest on its own.

Lemma 7. *Let F be a univariate distribution function with upper endpoint $x_F\in(0,\infty]$ such that F satisfies (3) with the positive scaling function w. Let further $a_n<b_n\le x_F,u_n,r_n,n\ge 1$ be four sequences of positive constants such that $u_n^*:=a_nu_n<x_F,\forall n\ge 1$*

$$\lim_{n\to\infty}u_n^*=x_F,\quad \textit{and}\quad \lim_{n\to\infty}u_nw(u_n^*)[b_n-a_n]=\eta\in[0,\infty]. \tag{38}$$

If further $\psi_n,h_n,n\ge1$ are positive measurable functions such that for all large n

$$\psi_n(a_n+x/(u_nw(u_n^*)))=r_nh_n(x),\quad \forall x>0, \tag{39}$$

where

$$\lim_{n\to\infty}h_n(x)\quad=\quad h(x),$$

and for all n large and any $x>0$

$$h_n(x)\quad\le\quad K\max(x^{\lambda_1},x^{\lambda_2}),\quad K\in(0,\infty),\lambda_i\in(-1,\infty),i=1,2,$$

is satisfied, then we have for any $\xi_n\to\xi\in[0,\infty)$ with $\xi\le\eta\le\infty$

$$\int_{a_n+\xi_n/(u_nw(u_n^*))}^{b_n}[1-F(u_nx)]\psi_n(x)\,dx\quad=\quad(1+o(1))\frac{r_n[1-F(u_n^*)]}{u_nw(u_n^*)}I(h,\eta,\xi)$$

as $n\to\infty$ with $I(h,\eta,\xi):=\int_\xi^\eta h(x)\exp(-x)\,ds\in[0,\infty)$.

Proof. Set for any $n\in\mathbb{N}$

$$u_n^*:=a_nu_n,\quad t_n:=u_nw(u_n^*),\quad \eta_n:=t_n[b_n-a_n].$$

Since $\lim_{n\to\infty}u_n^*=x_F$ the assumption on F implies

$$\lim_{n\to\infty}\frac{1-F(u_n^*+x/w(u_n^*))}{1-F(u_n^*)}=\exp(-x),\quad\forall x\in\mathbb{R}.$$

Next, applying Fatou Lemma we obtain

$$\begin{aligned}
&\liminf_{n\to\infty} \int_{a_n+\xi_n/w(u_n^*)}^{b_n} [1-F(u_n x)]\psi_n(x)\,dx \\
&\quad\geq \liminf_{n\to\infty} t_n^{-1} \int_{\xi_n}^{\eta_n} [1-F(u_n[a_n+x/t_n])]\psi_n(a_n+x/t_n)\,dx \\
&\quad\geq [1-F(u_n^*)] r_n t_n^{-1} \int_{\xi}^{\eta} \liminf_{n\to\infty} \frac{1-F(u_n^*+x/w(u_n^*)}{1-F(u_n^*)} h_n(x)\,dx \\
&\quad= (1+o(1)) r_n [1-F(u_n^*)] t_n^{-1} \int_{\xi}^{\eta} \exp(-x)h(x)\,dx, \quad n\to\infty.
\end{aligned}$$

Since for all x positive it follows that $h(x)\leq K\max(x^{\lambda_1},x^{\lambda_2})$ we have

$$0\leq \int_{\xi}^{\eta} \exp(-x)h(x)\,dx<\infty.$$

The proof for the lim sup follows (non-trivially) along the lines of the proof of Lemma 4.2, 4.3, 4.5 in Hashorva (2006a) utilising ideas and results in Berman (1992). The case $\lambda\geq 0$ follows easily with the arguments from Berman (1992) (see (12.3.7) therein). The case $\alpha\in(-1,0)$ is established using further the fact that for any $\varepsilon>0, x\in[0,1]$

$$\Big|\frac{1-F(\tau_n+x/w(\tau_n))}{1-F(\tau_n)}-\exp(-x)\Big|<\varepsilon\exp(-x)$$

holds uniformly for any sequence $\tau_n<x_F, n\geq 1$, such that $\lim_{n\to\infty}\tau_n=x_F$, hence the proof. □

Proof of Theorem 2 Set for $n\in\mathbb{N}$

$$a_n:=v_n/u_n,\quad w_n:=w(u_n),\quad q_n:=\sqrt{w_n/u_n}.$$

(6) implies $\lim_{n\to\infty}u_n w_n=\infty$. In view of (12) we have

$$q_n(Y_n-\rho_n u_n)|X_n>u_n \quad\stackrel{d}{\to}\quad Z_\rho,\quad n\to\infty,$$

with $Z_\rho/\sqrt{1-\rho^2}$ a standard Gaussian random variable. Consequently for all n large we have

$$\begin{aligned}
&\boldsymbol{P}\{X_n>u_n, Y_n>v_n+y_n/q_n\} \\
&\quad= \boldsymbol{P}\{X>u_n\}\boldsymbol{P}\{q_n(Y_n-\rho_n u_n)>q_n[a_n u_n+y_n/q_n-\rho_n u_n]|X_n>u_n\} \\
&\quad= \boldsymbol{P}\{X>u_n\}\boldsymbol{P}\{q_n(Y-\rho_n u_n)>y_n+\sqrt{u_n w_n}[a_n-\rho_n]|X>u_n\} \\
&\quad= (1+o(1))\boldsymbol{P}\{Z_\rho>y+z\}\boldsymbol{P}\{X>u_n\},\quad n\to\infty.
\end{aligned}$$

Using now (8) establishes the first claim.
ii) For simplicity assume that $\rho,\rho_n\in[0,1], n\geq 1$ and $v_n, n\geq 1$ is a positive sequence. The other case follows with similar arguments.
In view of Lemma 3.3 of Hashorva (2005b) we obtain

$$\begin{aligned}
&\boldsymbol{P}\{X_n>u_n, Y_n>a_n u_n\} \\
&\quad= \frac{1}{2\pi}\int_{\beta_n}^{\pi/2}[1-F(x/\cos(\alpha))]\,d\alpha+\frac{1}{2\pi}\int_{\psi_n-\pi/2}^{\beta_n}[1-F(y/\cos(\alpha-\psi_n))]\,d\alpha \\
&\quad=: I_{n1}+I_{n2},
\end{aligned}$$

with $\beta_n := \arctan((a_n-\rho_n)/\sqrt{1-\rho_n^2}), \psi_n := \arccos(\rho_n), n \geq 1$.
Define next for any $n \in \mathbb{N}$

$$\boldsymbol{\alpha}_{n,a,\rho} := 1/\cos(\beta_n) = \sqrt{(1-2\rho_n a_n + a_n^2)/(1-\rho_n^2)} \geq 1,$$

and

$$u_n^* := \boldsymbol{\alpha}_{n,a,\rho} u_n \quad w_n^* := w(u_n^*).$$

By the assumptions

$$\lim_{n\to\infty} \beta_n = \beta := \arctan((a-\rho)/\sqrt{1-\rho^2}),$$

and

$$\lim_{n\to\infty} \boldsymbol{\alpha}_{n,a,\rho} = 1/\cos(\beta) = \sqrt{(1-2\rho a + a^2)/(1-\rho^2)} > 1, \quad \lim_{n\to\infty} \boldsymbol{\alpha}_{n,a,\rho} u_n = \infty.$$

A simpler formula as the above one for the bivariate tail probability is given in Abdous et al. (2006), Klüppelberg et al. (2007). Transforming the variables (borrowing the idea and the formula from Abdous et al. (2006)) we obtain applying Lemma 7

$$\begin{aligned} I_{n1} &= \frac{1}{2\pi}\int_{1/\cos(\beta_n)}^{\infty} [1-F(u_n x)]\frac{1}{x}\frac{1}{\sqrt{x^2-1}}\,dx \\ &= (1+o(1))\frac{1-F(u_n^*)}{2\pi u_n w_n^*}\frac{1}{1/\cos(\beta)}\frac{1}{\sqrt{(1/\cos(\beta))^2-1}}\int_0^{\infty}\exp(-x)\,dx \\ &= (1+o(1))\frac{1-F(u_n^*)}{2\pi u_n w_n^*}\frac{1}{\boldsymbol{\alpha}_{a,\rho}}\frac{1}{\sqrt{(\boldsymbol{\alpha}_{a,\rho})^2-1}}, \quad n\to\infty, \end{aligned}$$

and similarly for any $a > 0$

$$\begin{aligned} I_{n2} &= \frac{1}{2\pi}\int_{a_n/\cos(\beta_n)}^{\infty} [1-F(a_n u_n y)]\frac{1}{y}\frac{1}{\sqrt{y^2-1}}\,dy \\ &= (1+o(1))\frac{1-F(u_n^*)}{2\pi u_n w_n^*}\frac{1}{\boldsymbol{\alpha}_{a,\rho}}\frac{1}{\sqrt{(\boldsymbol{\alpha}_{a,\rho})^2/a^2-1}}, \quad n\to\infty. \end{aligned}$$

Consequently we may write as $n\to\infty$ using further (8)

$$\begin{aligned} &\boldsymbol{P}\{X_n > u_n, Y_n > a_n u_n\} \\ &= (1+o(1))\frac{1}{\boldsymbol{\alpha}_{a,\rho}}\Big[\frac{1}{\sqrt{(\boldsymbol{\alpha}_{a,\rho})^2-1}} + \frac{1}{\sqrt{(\boldsymbol{\alpha}_{a,\rho})^2/a^2-1}}\Big]\frac{1-F(u_n^*)}{2\pi u_n w_n^*} \\ &= (1+o(1))\frac{\boldsymbol{\alpha}_{a,\rho}C_{a,\rho}}{2\pi}\frac{1-F(u_n^*)}{u_n w_n^*} \\ &= (1+o(1))\frac{\boldsymbol{\alpha}_{a,\rho}^2 C_{a,\rho}}{\sqrt{2\pi}}\left(\frac{1}{u_n^* w_n^*}\right)^{1/2}\frac{1-F(u_n^*)}{\sqrt{2\pi\boldsymbol{\alpha}_{n,a,\rho}u_n w_n^*}} \\ &= (1+o(1))\frac{\boldsymbol{\alpha}_{a,\rho}^2 C_{a,\rho}}{\sqrt{2\pi}}\left(\frac{1}{u_n^* w_n^*}\right)^{1/2}\boldsymbol{P}\{X > u_n^*\}, \quad n\to\infty. \end{aligned}$$

In the case $a = 0$ we obtain the same asymptotics since $I_{n2} = o(I_{n1}), n\to\infty$. Thus the claim follows. □

Proof of Corollary 3 In view of (21)

$$\lim_{n\to\infty}\sqrt{u_n w(u_n)}[a-\rho] = 0, \quad \text{if } a = \rho$$

or

$$\lim_{n\to\infty}\sqrt{u_n w(u_n)}[a-\rho]=-\infty,\quad \text{if } a<\rho$$

implying that (14) holds with $z=0$ or $z=-\infty$, respectively. Hence the proof follows immediately from Theorem 2. □

PROOF OF COROLLARY 4 i) Since $u_n\to\infty$ we get using (6)

$$\lim_{n\to\infty}\sqrt{u_n w(u_n)}[y/u_n-\rho]=-\infty,$$

hence (14) holds with $z=-\infty$. Applying Theorem 2 establishes the first claim.
ii) In this case (14) holds with $z=by$, thus the claim follows again by a direct application of the mentioned theorem.
iii) For all large n we have $a_n:=y/u_n>\rho$ and further $\lim_{n\to\infty}a_n=0$. Utilising again Theorem 2 establishes the proof. □

PROOF OF THEOREM 5 (5) and (8) imply that both S_1 and S_2 are in the Gumbel max-domain of attraction with the scaling function w. Consequently

$$\lim_{n\to\infty}\frac{\boldsymbol{P}\{S_1>u_n+xr_n/w(u_n)\}}{\boldsymbol{P}\{S_1>u_n\}}=0$$

for any $r_n, n\geq 1$ tending to ∞. Hence the proof follows easily from (12). □

PROOF OF THEOREM 6 Set for $n\geq 1$

$$q_n:=\sqrt{w(u_n)/u_n},\quad u_n':=u_n+h_{n1}x,\quad y_n=yq_nh_{n2}.$$

By the assumptions and using (6) we obtain for any $x,y\in\mathbb{R}$

$$\lim_{n\to\infty}u_n'=\infty,\quad \lim_{n\to\infty}y_n=c_2y$$

and

$$\lim_{n\to\infty}\frac{u_n'}{u_n}=\lim_{n\to\infty}[1+xh_{n1}/u_n]=\lim_{n\to\infty}[1+x\frac{(1+o(1))c_1}{\sqrt{u_nw(u_n)}}]=1,$$

hence

$$\begin{aligned}\lim_{n\to\infty}\left(\frac{w(u_n')}{u_n'}\right)^{1/2}[v_n-\rho_nu_n'] &= \lim_{n\to\infty}q_n[v_n-\rho_nu_n]-\rho_nx\lim_{n\to\infty}h_{n1}q_n\\ &= z-\rho c_1x,\end{aligned}$$

consequently applying Theorem 2 we obtain for any $x,y\in\mathbb{R}$

$$\begin{aligned}&\boldsymbol{P}\{X_n>u_n+h_{n1}x,Y_n>v_n+h_{n2}y\}\\ &=\boldsymbol{P}\{X_n>u_n',Y_n>v_n+y_n\sqrt{u_n/w(u_n)}\}\\ &=(1+o(1))\boldsymbol{P}\{Z_\rho>c_2y-\rho c_1x+z\}\boldsymbol{P}\{X>u_n+h_{n1}x\},\quad n\to\infty.\end{aligned}$$

Thus we have if x,y are positive (recall (7))

$$\begin{aligned}&\boldsymbol{P}\{X_n>u_n+h_{n1}x,Y_n>v_n+h_{n2}y\big|X_n>u_n,Y_n>v_n\}\\ &=\frac{\boldsymbol{P}\{X_n>u_n+h_{n1}x,Y_n>v_n+h_{n2}y\}}{\boldsymbol{P}\{X_n>u_n,Y_n>v_n\}}\\ &=(1+o(1))\overline{\Phi}_{\rho,z}(c_2y-\rho c_1x)\frac{\boldsymbol{P}\{X>u_n+h_{n1}x\}}{\boldsymbol{P}\{X>u_n\}},\quad n\to\infty,\end{aligned}$$

with $\overline{\Phi}_{\rho,z}:=\boldsymbol{P}\{Z_\rho>s+z\}/\boldsymbol{P}\{Z_\rho>z\}, s\in\mathbb{R}$ and $Z_\rho/\sqrt{1-\rho^2}$ a standard Gaussian random variable, hence (34) follows.

Next, if the sequence $h_{n1}, n \geq 1$ is asymptoticaly equivalent with $w(u_n), n \geq 1$, i.e., $\lim_{n\to\infty} h_{n1} w(u_n) = 1$, then

$$\lim_{n\to\infty} q_n h_{n1} = 0,$$

and further $\forall x \in \mathbb{R}$ (recall (8))

$$\lim_{n\to\infty} \frac{1 - F(u_n + h_{n1}x)}{1 - F(u_n)} = \lim_{n\to\infty} \frac{\boldsymbol{P}\{X > u_n + x/w(u_n)\}}{\boldsymbol{P}\{X > u_n\}} = \exp(-x).$$

Consequently if additionally $\lim_{n\to\infty} h_{n2} w(u_n) = 1$ we obtain for any $x, y \in [0, \infty)$

$$\lim_{n\to\infty} \boldsymbol{P}\{X_n > u_n + h_{n1}x, Y_n > v_n + h_{n2}y \big| X_n > u_n, Y_n > v_n\} = \exp(-x)\overline{\Phi}_{\rho,z}(y).$$

We thus have the convergence in distribution

$$w(u_n)(X_n - u_n)\big|X_n > u_n, Y_n > v_n \quad \stackrel{d}{\to} \quad U, \quad n \to \infty,$$

with U a unit exponential random variable, and for any $z \in \mathbb{R}$

$$\sqrt{w(u_n)/u_n}(Y_n - v_n)\big|X_n > u_n, \quad Y_n > v_n \quad \stackrel{d}{\to} \quad V_z, \quad n \to \infty,$$

where V_z is a positive random variable with survival function $\overline{\Phi}_{\rho,z}(y), y \geq 0$. Furthermore, the joint convergence in distribution holds, hence (34) follows.
ii) Since $\lim_{n\to\infty} u_n w(u_n) = \infty$ we have that (14) holds with $z = \infty$, hence Theorem 2 implies for any $x, y \in \mathbb{R}$ as $n \to \infty$

$$\begin{aligned} &\boldsymbol{P}\{X_n > u_n + h_{n1}y, Y_n > v_n + h_{n2}y\} \\ &= \quad (1 + o(1))\frac{\boldsymbol{\alpha}_{a,\rho} C_{a,\rho}}{2\pi}\left(\frac{1}{u_n w(t_n^*)}\right)[1 - F(t_n^*)], \end{aligned}$$

with $C_{a,\rho}$ defined in (9) and

$$t_n^* = u_n^* + (1 + o(1))[(\boldsymbol{\alpha}_{a,\rho} - aK_{a,\rho})x + K_{a,\rho}y]/w(u_n^*), \quad n \to \infty,$$

where $u_n^* := \boldsymbol{\alpha}_{n,a,\rho} u_n, \boldsymbol{\alpha}_{n,a,\rho} := \sqrt{(1 - 2\rho_n a_n + a_n^2)/(1 - \rho_n^2)}$, and

$$\boldsymbol{\alpha}_{a,\rho} := \sqrt{(1 - 2a\rho + a^2)/(1 - \rho^2)} > 1, \quad K_{a,\rho} := \frac{a - \rho}{\boldsymbol{\alpha}_{a,\rho}(1 - \rho^2)} > 0.$$

Since

$$\boldsymbol{\alpha}_{a,\rho} - aK_{a,\rho} = \frac{1 - a\rho}{\boldsymbol{\alpha}_{a,\rho}(1 - \rho^2)} =: \overline{K}_{a,\rho} > 0$$

we may further write ($n \to \infty$)

$$\begin{aligned} &\boldsymbol{P}\Big\{X_n > u_n + \frac{x}{w(u_n^*)}, Y_n > v_n + \frac{y}{w(u_n^*)}\Big\} \\ &= \quad (1 + o(1))\frac{\boldsymbol{\alpha}_{a,\rho} C_{a,\rho}}{2\pi} \exp(-\overline{K}_{a,\rho}x - K_{a,\rho}y)\left(\frac{1}{u_n w(u_n^*)}\right)^{1/2}[1 - F(u_n^*)], \end{aligned}$$

thus the proof follows. □

References

[1] Abdous, B., Fougères, A.-L., and Ghoudi, K. (2005) Extreme behaviour for bivariate elliptical distributions. *The Canadian Journal of Statistics* **33**,(3), 317-334.

[2] Abdous, B., Fougères, A.-L., Ghoudi, K., and Soulier, P. (2006) Estimation of bivariate excess probabilities for elliptical models. (*www.arXiv:math.ST/0611914*).

[3] Anderson, T.W., and Fang, K.T. (1990) On the theory of multivariate elliptically contoured distributions and their applications. In *Statistical Inference in Elliptically Contoured and Related Distributions*, K.T. Fang and T.W. Anderson, eds, Allerton Press, New York, pp. 1–23.

[4] Asimit, A.V., and Jones, B.L. (2007) Extreme behavior of bivariate elliptical distributions. *Insurance: Mathematics and Economics*, **41**, 1, 53–61.

[5] Berman, M.S. (1962) A low of large numbers for the maximum in a stationary Gaussian sequence. *Ann. Math. Stats.* **33**, (1), 93–97.

[6] Berman, M.S. (1982) Sojourns and extremes of stationary processes. *Ann. Probability* **10**, 1–46.

[7] Berman, M.S. (1983) Sojourns and extremes of Fourier sums and series with random coefficients. *Stoch. Proc. Appl.* **15**, 213–238.

[8] Berman, M.S. (1992) *Sojourns and Extremes of Stochastic Processes*. Wadsworth & Brooks/ Cole, Boston.

[9] Butler, A., and Tawn, J.A. (2005) Conditional extremes of a markov chain. Preprint.

[10] Cambanis, S., Huang, S., and Simons, G. (1981) On the theory of elliptically contuored distributions. *J. Multivariate Anal.* **11**, 368–385.

[11] Carnal, H. (1970) Die konvexe Hülle von n rotations-symmetrisch verteilten Punkten. *Z. Wahrscheinlichkeitstheorie Verw. Geb.* **15**, 168-176.

[12] Dai, M., and Mukherjea, A. (2001) Identification of the parameters of a multivariate normal vector by the distribution of the minimum. *J. Theoretical Prob.* **14**, 1, 267–298.

[13] De Haan, L. (1970) *On Regular Variation and its Applications to the Weak Convergence of Sample Extremes.* Mathematisch Centrum Amsterdam, Netherlands.

[14] De Haan, L., and Ferreira, A. (2006) *Extreme Value Theory. An Introduction*. Springer.

[15] Eddy, W.F., and Gale, J.D. (1981) The convex hull of a spherically symmetric sample. *Advances in Applied Probability*, **13**, 751–763.

[16] Falk, M., Hüsler, J., and Reiss R.-D. (2004) *Laws of Small Numbers: Extremes and Rare Events.* DMV Seminar **23**, 2-nd edition, Birkhäuser, Basel.

[17] Fang, K.-T., Kotz, S., and Ng, K.-W. (1990) *Symmetric Multivariate and Related Distributions.* Chapman and Hall, London, United Kingdom.

[18] Fang, K., and Zhang, Y. (1990) *Generalized Multivariate Analysis.* Springer, Berlin, Heidelberg, New York.

[19] Galambos, J. (1987) *Asymptotic Theory of Extreme Order Statistics*, 2nd ed. Krieger, Malabar, Florida.

[20] Gale, J.D. (1980) The Asymptotic Distribution of the Convex Hull of a Random Sample. *Ph.D. Thesis, Carnegie-Mellon University.*

[21] Gupta, A.K., and Varga, T. (1993) *Elliptically Contoured Models in Statistics.* Kluwer, Dordrecht.

[22] Hashorva, E. (2005a) Asymptotics and bounds for multivariate Gaussian tails. *J. Theoretical Prob.* **18**, 1,79-97.

[23] Hashorva, E. (2005b) Elliptical triangular arrays in the max-domain of attraction of Hüsler-Reiss distributon. *Statist. Probab. Lett.* **72** (2), 125–135.

[24] Hashorva, E. (2006a) Gaussian approximation of conditional elliptical random vectors. *Stochastic Models.* **22**, 441-457.

[25] Hashorva, E. (2006b) Exact asymptotic behaviour of Type I elliptical random vectors. Submitted. (*http://www.imsv.unibe.ch/~ enkelejd/mextrtypI.pdf*)

[26] Hashorva, E., Kotz, S., and Kume, A. (2007) L_p-norm generalised symmetrised Dirichlet distributions. *Albanian Journal of Mathematics.* **1** (1), 31–56.

[27] Hashorva, E., and Hüsler J. (2003) On Multivariate Gaussian Tails. *Ann. Inst. Statist. Math.* **55**,(3), 507–522.

[28] Heffernan, J.E., and Tawn, J.A. (2004) A conditional approach for multivariate extreme values. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **66**, 3, 497-546.

[29] Heffernan, J.E., and Resnick, S.I. (2005) Limit laws for random vectors with an extreme component. (*http://www.maths.lancs.ac.uk/~currie/Papers/ConditModel.pdf*)
[30] Kano, Y. (1994) Consistency property of elliptical probability density functions. *J. Multivariate Anal.* 51, 139–147.
[31] Klüppelberg, C., Kuhn, K., and Peng, L. (2007) Estimating the tail dependence of an elliptical distribution. *Bernoulli*, **13** (1), 229–251.
[32] Kotz, S. (1975) Multivariate distributions at a cross-road. In: *Statistical Distributions in Scientific Work* 1, G.P. Patil, S. Kotz, and J.K. Ordeds, D. Riedel, Dortrecht, 240–247.
[33] Kotz, S., and Ostrovskii, I.V. (1994) Characteristic functions of a class of elliptical distributions. *J. Multivariate Analysis*. **49**, (1), 164–178.
[34] Kotz, S., and Nadarajah, S. (2005) *Extreme Value Distributions, Theory and Applications*. Imperial College Press, London, United Kingdom. (Second Printing).
[35] Leadbetter, M.R., Lindgren, G., and Rootzén, H. (1983) *Extremes and related properties of random sequences and processes*. Springer-Verlag, New York.
[36] Reiss, R-D. (1989) *Approximate Distributions of Order Statistics: With Applications to Nonparametric Statistics*. Springer, New York.
[37] Resnick, S.I. (1987) *Extreme Values, Regular Variation and Point Processes*. Springer, New York.
[38] Szabłowski, P.L. (1990) Expansions of $E(X|Y+\epsilon X)$ and their applications to the analysis of elliptically contoured measures. *Comput. Math. Appl.* 19, No.5, 75–83.

University of Bern, Institute of Mathematical Statistics and Actuarial Sciences,, Sidlerstrasse 5, CH-3012 Bern, Switzerland, and, Allianz Suisse Insurance Company, Laupenstrasse 27, CH-3001 Bern, Switzerland

E-mail address: enkelejd.hashorva@stat.unibe.ch

E-mail address: enkelejd.hashorva@Allianz-Suisse.ch

ALBANIAN JOURNAL
OF MATHEMATICS
Volume 1, Number 2, Pages 115–120
ISSN 1930-1235: (electronic version)

MAPPINGS AND PAIRWISE CONTINUITY ON PAIRWISE LINDELÖF BITOPOLOGICAL SPACES

ADEM KILIÇMAN AND ZABIDIN SALLEH

Abstract. In this paper, we shall continue the study of pairwise Lindelöf bitopological spaces initiated by Fora and Hdeib. Furthermore, we introduce the concepts of pairwise continuity, pairwise open and pairwise homeomorphism, and the properties for them are established. We also shows that a Lindelöf space and a B-Lindelöf space are both bitopological property and p-topological property.

1. Introduction

A bitopological space (X, τ_1, τ_2) is a set X together with two topologies τ_1 and τ_2 defined on X. The reader is suggested to refer to [3] for the detail definitions, terminology and notation.

Nowadays mapping and continuity stand among the most important topics and most researched points in topology. It is being studied by many general topologists including the authors. In this paper we extend the idea of continuity in a topological space to a bitopological space. We also extend the result for the continuity in a topological space to the pairwise continuity in a bitopological space and study their properties.

The purpose of this paper is to study the effect of mappings and pairwise continuity on pairwise Lindelöf bitopological spaces. These spaces were introduced by Fora and Hdeib [2]. We show that some mappings preserve these properties (i.e., Lindelöf property and B-Lindelöf property). The main results in our study are that the image of a Lindelöf bitopological space under a continuous or a p-continuous function is Lindelöf, and the image of a B-Lindelöf space under a continuous and open function or a p-continuous and p-open function is B-Lindelöf.

2. Preliminaries

Throughout this paper, all spaces (X, τ) and (X, τ_1, τ_2) (or simply X) always mean topological spaces and bitopological spaces, respectively. In this paper, we shall use p- to denote pairwise. For instance, p-continuous stands for pairwise continuous. We always use (τ_i, τ_j)- to denote certain properties with respect to topology τ_i and τ_j in bitopological spaces, where $i, j \in \{1, 2\}$. By τ_i-open set, we shall mean the open set with respect to topology τ_i in X. By τ_i-open cover of X, we mean that the cover of X by τ_i-open sets in X. Sometimes the prefixes (τ_i, τ_j)- or τ_i- will

2000 *Mathematics Subject Classification.* 54E55.
Key words and phrases. Bitopological spaces, Lindelöf, B-Lindelöf , continuous, p-continuous, open, p-open, homeomorphism, p-homeomorphism.

be replaced by (i,j)- or i-, respectively, if there is no chance for confusion. The reader may consult [1] for the detail notations. The authors sometime write the term "pairwise Lindelöf spaces" meaning pairwise Lindelöf bitopological spaces.

Definition 1 (see [2, 4]). *A bitopological space (X,τ_1,τ_2) is said to be Lindelöf if the topological space (X,τ_1) and (X,τ_2) are both Lindelöf. Equivalently, (X,τ_1,τ_2) is i-Lindelöf if the topological space (X,τ_i) is Lindelöf. X is said Lindelöf if it is i-Lindelöf for each $i=1,2$, or, if every i-open cover of X has a countable subcover for each $i=1,2$.*

Definition 2 (see [2]). *A bitopological space (X,τ_1,τ_2) is called (i,j)-Lindelöf if for every i-open cover of X there is a countable j-open subcover. X is called B-Lindelöf if it is both (i,j)-Lindelöf and (j,i)-Lindelöf.*

Kopperman [5] and Tallafha et. al. [6], was mentioned about pairwise continuous functions and pairwise open functions on bitopological spaces. Tallafha et. al. [6] has given more definitions, i.e., pairwise closed and pairwise homeomorphism functions. The following assert the definition of continuous function in the sense of Tallafha et. al.

Definition 3. *Let (X,τ_1,τ_2) and (Y,σ_1,σ_2) be two bitopological spaces. A function $f:(X,\tau_1,\tau_2)\to(Y,\sigma_1,\sigma_2)$ is said to be continuous if the functions $f:(X,\tau_1)\to(Y,\sigma_1)$ and $f:(X,\tau_2)\to(Y,\sigma_2)$ are both continuous. Equivalently, a function $f:(X,\tau_1,\tau_2)\to(Y,\sigma_1,\sigma_2)$ is called i-continuous if the function $f:(X,\tau_i)\to(Y,\sigma_i)$ is continuous. f is said continuous if it is i-continuous for each $i=1,2$.*

Next we are going to define the second concept of pairwise continuous function on bitopological spaces in the sense of Tallafha et. al.

Definition 4. *Let (X,τ_1,τ_2) and (Y,σ_1,σ_2) be two bitopological spaces. A function $f:(X,\tau_1,\tau_2)\to(Y,\sigma_1,\sigma_2)$ is called (i,j)-continuous if the function $f:(X,\tau_i)\to(Y,\sigma_j)$ is continuous. The function $f:(X,\tau_1,\tau_2)\to(Y,\sigma_1,\sigma_2)$ is called p-continuous if it is both (i,j)-continuous and (j,i)-continuous. Equivalently, a function $f:(X,\tau_1,\tau_2)\to(Y,\sigma_1,\sigma_2)$ is p-continuous if the functions $f:(X,\tau_1)\to(Y,\sigma_2)$ and $f:(X,\tau_2)\to(Y,\sigma_1)$ are both continuous.*

The following definitions are given two concepts of pairwise open and pairwise closed functions in the sense of Tallafha et. al.

Definition 5. *A function $f:(X,\tau_1,\tau_2)\to(Y,\sigma_1,\sigma_2)$ is said to be open (resp. closed) if the functions $f:(X,\tau_1)\to(Y,\sigma_1)$ and $f:(X,\tau_2)\to(Y,\sigma_2)$ are both open (resp. closed). Equivalently, a function $f:(X,\tau_1,\tau_2)\to(Y,\sigma_1,\sigma_2)$ is called i-open (resp. i-closed) if the function $f:(X,\tau_i)\to(Y,\sigma_i)$ is open (resp. closed). f is said open (resp. closed) if f is i-open (resp. i-closed) for each $i=1,2$.*

Definition 6. *A function $f:(X,\tau_1,\tau_2)\to(Y,\sigma_1,\sigma_2)$ is called (i,j)-open (resp. (i,j)-closed) if the function $f:(X,\tau_i)\to(Y,\sigma_j)$ is open (resp. closed). f is said p-open (resp. p-closed) if it is both (i,j)-open (resp. (i,j)-closed) and (j,i)-open (resp. (j,i)-closed). Equivalently, a function $f:(X,\tau_1,\tau_2)\to(Y,\sigma_1,\sigma_2)$ is said to be p-open (resp. p-closed) if the functions $f:(X,\tau_1)\to(Y,\sigma_2)$ and $f:(X,\tau_2)\to(Y,\sigma_1)$ are both open (resp. closed).*

Example 1. *Consider $X=\{a,b,c,d\}$ with τ_1 the discrete topology and topology $\tau_2=\{\emptyset,\{a\},\{a,b\},\{a,b,c\},X\}$ on X, and $Y=\{x,y,z,w\}$ with topologies $\sigma_1=$*

$\{\emptyset, \{x\}, \{y\}, \{x, y\}, \{y, z, w\}, Y\}$ *and* $\sigma_2 = \{\emptyset, \{x\}, \{y, z, w\}, Y\}$ *on* Y. *Define a function* $f : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ *by* $f(a) = y, f(b) = f(d) = z$ *and* $f(c) = w$. *Observe that the functions* $f : (X, \tau_1) \to (Y, \sigma_1)$ *and* $f : (X, \tau_2) \to (Y, \sigma_2)$ *are continuous. Therefore the function* $f : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ *is continuous. On the other hand, the functions* $f : (X, \tau_1) \to (Y, \sigma_2)$ *and* $f : (X, \tau_2) \to (Y, \sigma_1)$ *are also continuous. Therefore the function* $f : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ *is p-continuous.*

Example 2. *Consider* (X, τ_1, τ_2) *and* (Y, σ_1, σ_2) *are bitopological spaces as in Example* 1. *Define a function* $g : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ *by* $g(a) = g(b) = x, g(c) = z$ *and* $g(d) = w$. *The function* $g : (X, \tau_1) \to (Y, \sigma_1)$ *is continuous and* $g : (X, \tau_2) \to (Y, \sigma_2)$ *is not continuous since* $\{y, z, w\} \in \sigma_2$ *but its inverse image* $g^{-1}(\{y, z, w\}) = \{c, d\} \notin \tau_2$. *Thus* $g : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ *is not continuous. On the other hand, the function* $g : (X, \tau_1) \to (Y, \sigma_2)$ *is continuous and* $g : (X, \tau_2) \to (Y, \sigma_1)$ *is not continuous since* $\{y, z, w\} \in \sigma_1$ *but* $g^{-1}(\{y, z, w\}) = \{c, d\} \notin \tau_2$. *Thus* $g : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ *is not p-continuous.*

Example 3. *Consider a function* $f : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ *as in Example* 1. *Observe that the function* $f : (X, \tau_2) \to (Y, \sigma_2)$ *is not open since* $\{a\} \in \tau_2$ *but* $f(\{a\}) = \{y\} \notin \sigma_2$. *Thus* $f : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ *is not open. The function* $f : (X, \tau_2) \to (Y, \sigma_1)$ *is not open since* $\{a, b\} \in \tau_2$ *but* $f(\{a, b\}) = \{y, z\} \notin \sigma_1$. *Thus* $f : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ *is not p-open.*

3. Some Results on Pairwise Continuity

In this section, we are going to study the images for each type of pairwise Lindelöf spaces under several types of combinations of pairwise continuous and pairwise open functions. We shows that some mappings preserve certain type of pairwise Lindelöf spaces and the other, the images are another type of pairwise Lindelöf spaces.

Theorem 1. *Let* $f : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ *be an* i*-continuous (resp.* (j, i)*-continuous) and surjective function. If* (X, τ_1, τ_2) *is* τ_i*-Lindelöf (resp.* τ_j*-Lindelöf), then* (Y, σ_1, σ_2) *is* σ_i*-Lindelöf.*

Proof. Let $\{G_k : k \in \Delta\}$ be a σ_i-open cover of Y, i.e., $Y = \bigcup_{k \in \Delta} G_k$ where $G_k \in \sigma_i$. Since $f : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ is i-continuous (resp. (j, i)-continuous), then $f^{-1}(G_k) \in \tau_i$ (resp. $f^{-1}(G_k) \in \tau_j$) and $X = f^{-1}(Y) = \bigcup_{k \in \Delta} f^{-1}(G_k)$. Hence $\{f^{-1}(G_k) : k \in \Delta\}$ is a τ_i-open (resp. τ_j-open) cover of X. Since (X, τ_1, τ_2) is τ_i-Lindelöf (resp. τ_j-Lindelöf), so there exists a countable τ_i-open (resp. τ_j-open) subcover of X, say $\{f^{-1}(G_{k_n}) : n \in \mathbb{N}\}$ such that $X = \bigcup_{n \in \mathbb{N}} f^{-1} G_{k_n}$. Since f is surjective, $Y = f(X) = \bigcup_{n \in \mathbb{N}} f(f^{-1}(G_{k_n})) \subseteq \bigcup_{n \in \mathbb{N}} G_{k_n}$. Thus we obtain $\{G_{k_n} : n \in \mathbb{N}\}$ is a countable σ_i-open subcover of Y. Therefore (Y, σ_1, σ_2) is σ_i-Lindelöf. □

Corollary 1. *Let* $f : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ *be a continuous (resp. p-continuous) and surjective function. If* (X, τ_1, τ_2) *is Lindelöf, then* (Y, σ_1, σ_2) *is Lindelöf.*

Theorem 2. *Let* $f : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ *be an* i*-continuous (resp.* (i, j)*-continuous), surjective and* j*-open (resp.* (j, i)*-open) function. If* (X, τ_1, τ_2) *is* (τ_i, τ_j)*-Lindelöf, then* (Y, σ_1, σ_2) *is* (σ_i, σ_j)*-Lindelöf (resp.* (σ_j, σ_i)*-Lindelöf).*

Proof. Let $\{G_k : k \in \Delta\}$ is a σ_i-open (resp. σ_j-open) cover of Y. Following the proof of Theorem 1, $\left\{f^{-1}(G_k) : k \in \Delta\right\}$ is a τ_i-open cover of X. Since (X, τ_1, τ_2) is (τ_i, τ_j)-Lindelöf, so the τ_i-open cover of X has a countable τ_j-open subcover, say $\left\{f^{-1}(G_{k_n}) : n \in \mathbb{N}\right\}$ such that $X = \bigcup_{n\in\mathbb{N}} f^{-1}(G_{k_n})$. Since f is surjective, $Y = f(X) = \bigcup_{n\in\mathbb{N}} f\left(f^{-1}(G_{k_n})\right) \subseteq \bigcup_{n\in\mathbb{N}} G_{k_n}$. Thus we obtain $\{G_{k_n} : n \in \mathbb{N}\}$ is a countable subfamily by σ_j-open (resp. σ_i-open) sets which cover Y since f is a j-open (resp. (j,i)-open) function. Therefore (Y, σ_1, σ_2) is (σ_i, σ_j)-Lindelöf (resp. (σ_j, σ_i)-Lindelöf). □

Corollary 2. *Let $f : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ be a continuous (resp. p-continuous), surjective and open (resp. p-open) function. If (X, τ_1, τ_2) is B-Lindelöf, then (Y, σ_1, σ_2) is B-Lindelöf.*

Theorem 3. *Let $f : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ be an i-continuous (resp. (j,i)-continuous), surjective and (i,j)-open (resp. j-open) function. If (X, τ_1, τ_2) is τ_i-Lindelöf (resp. τ_j-Lindelöf), then (Y, σ_1, σ_2) is (σ_i, σ_j)-Lindelöf.*

Proof. Let $\{G_k : k \in \Delta\}$ be a σ_i-open cover of Y. Since $f : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ is i-continuous (resp. (j,i)-continuous), then $\left\{f^{-1}(G_k) : k \in \Delta\right\}$ is a τ_i-open (resp. τ_j-open) cover of X. Since (X, τ_1, τ_2) is τ_i-Lindelöf (resp. τ_j-Lindelöf), so there exists a countable τ_i-open (resp. τ_j-open) subcover of X, say $\left\{f^{-1}(G_{k_n}) : n \in \mathbb{N}\right\}$. Since f is surjective and (i,j)-open (resp. j-open) function, we obtain $\{G_{k_n} : n \in \mathbb{N}\}$ is a countable subfamily by σ_j-open sets which also cover Y. This shows that (Y, σ_1, σ_2) is (σ_i, σ_j)-Lindelöf. □

Corollary 3. *Let $f : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ be a continuous (resp. p-continuous), surjective and p-open (resp. open) function. If (X, τ_1, τ_2) is Lindelöf, then (Y, σ_1, σ_2) is B-Lindelöf.*

Theorem 4. *Let $f : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ be an i-continuous (resp. (j,i)-continuous) and surjective function. If (X, τ_1, τ_2) is (τ_i, τ_j)-Lindelöf (resp. (τ_j, τ_i)-Lindelöf), then (Y, σ_1, σ_2) is σ_i-Lindelöf.*

Proof. Let $\{G_k : k \in \Delta\}$ be a σ_i-open cover of Y. Since $f : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ is i-continuous (resp. (j,i)-continuous), then $\left\{f^{-1}(G_k) : k \in \Delta\right\}$ is a τ_i-open (resp. τ_j-open) cover of X. Since (X, τ_1, τ_2) is (τ_i, τ_j)-Lindelöf (resp. (τ_j, τ_i)-Lindelöf), so the τ_i-open (resp. τ_j-open) cover of X has a countable τ_j-open (resp. τ_i-open) subcover, say $\left\{f^{-1}(G_{k_n}) : n \in \mathbb{N}\right\}$. Since f is surjective, we obtain $\{G_{k_n} : n \in \mathbb{N}\}$ is a countable σ_i-open subcover of Y. This shows that (Y, σ_1, σ_2) is σ_i-Lindelöf. □

Corollary 4. *Let $f : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ be a continuous (resp. p-continuous) and surjective function. If (X, τ_1, τ_2) is B-Lindelöf, then (Y, σ_1, σ_2) is Lindelöf.*

The concept of homeomorphism is well known in topological spaces. Now we extend this concept to bitopological spaces in the following definitions in sense of Tallafha et. al. [6].

Definition 7. *Let (X, τ_1, τ_2) and (Y, σ_1, σ_2) be two bitopological spaces. Then a function $f : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ is called i-homeomorphism if the function $f : (X, \tau_i) \to (Y, \sigma_i)$ is homeomorphism, or equivalently, if $f : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$*

is bijection, i-continuous and $f^{-1} : (Y, \sigma_1, \sigma_2) \to (X, \tau_1, \tau_2)$ *is i-continuous. The bitopological spaces* (X, τ_1, τ_2) *and* (Y, σ_1, σ_2) *are then called i-homeomorphic.*

A function $f : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ *is called homeomorphism if the function* $f : (X, \tau_i) \to (Y, \sigma_i)$ *is homeomorphism for each* $i = 1, 2$*, or equivalently, if* $f : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ *is bijection, continuous and* $f^{-1} : (Y, \sigma_1, \sigma_2) \to (X, \tau_1, \tau_2)$ *is continuous. The bitopological spaces* (X, τ_1, τ_2) *and* (Y, σ_1, σ_2) *are then called homeomorphic.*

In the following definition is given the second type of pairwise homeomorphism functions in the sense of Tallafha et. al. [6].

Definition 8. *Let* (X, τ_1, τ_2) *and* (Y, σ_1, σ_2) *be two bitopological spaces. Then a function* $f : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ *is called* (i, j)*-homeomorphism if the functions* $f : (X, \tau_i) \to (Y, \sigma_j)$ *is homeomorphism, or equivalently, if* $f : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ *is bijection,* (i, j)*-continuous and* $f^{-1} : (Y, \sigma_1, \sigma_2) \to (X, \tau_1, \tau_2)$ *is* (i, j)*-continuous. The bitopological spaces* (X, τ_1, τ_2) *and* (Y, σ_1, σ_2) *are then called* (i, j)*-homeomorphic.*

A function $f : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ *is called p-homeomorphism if the function f is both* (i, j)*-homeomorphism and* (j, i)*-homeomorphism, or equivalently, if* $f : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ *is bijection, p-continuous and* $f^{-1} : (Y, \sigma_1, \sigma_2) \to (X, \tau_1, \tau_2)$ *is p-continuous. The bitopological spaces* (X, τ_1, τ_2) *and* (Y, σ_1, σ_2) *are then called p-homeomorphic.*

Tallafha et. al. [6] use notation p_1- instead of p- for our definition of second type of pairwise. While Fora and Hdeib [2], and Kiliçman and Salleh [4] use notation p- instead of our definition of first type of pairwise. For instance Fora and Hdeib use p-continuous instead of continuous.

The concept of p- that was defined throughout this paper depends heavily on the order of the topologies, that is, (X, τ_1, τ_2) is different from (X, τ_2, τ_1). For instance, $(\mathbb{R}, \tau_u, \tau_{cof})$ is p-homeomorphic to $(\mathbb{R}, \tau_{cof}, \tau_u)$ but it is not p-homeomorphic to itself, where τ_u and τ_{cof} are usual topology and cofinite topology on $\mathbb{R}$, respectively. However, a bitopological space (X, τ_1, τ_2) is p-homeomorphic to itself if and only if (X, τ_1) is homeomorphic to (X, τ_2). Similar for the p-open, p-closed and p-continuous functions.

Example 4. *The function* $f : (X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ *in Example* 1 *is not homeomorphism and not p-homeomorphism since* $f^{-1} : (Y, \sigma_1, \sigma_2) \to (X, \tau_1, \tau_2)$ *is not continuous and not p-continuous.*

Recall that, a property $\mathcal{P}$ of sets is called topological property if whenever a topological space (X, τ) has property $\mathcal{P}$, then every space homeomorphic to (X, τ) also has property $\mathcal{P}$. In the case of bitopological space (X, τ_1, τ_2), there are two types of topological properties since we have two types of homeomorphism. A property $\mathcal{P}$ will be called i-topological property (resp. (i, j)-topological property) if whenever (X, τ_1, τ_2) has property $\mathcal{P}$, then every space i-homeomorphic (resp. (i, j)-homeomorphic) to (X, τ_1, τ_2) also has property $\mathcal{P}$. If homeomorphism (resp. p-homeomorphism) considered (for the pairwise topology), we will call such property $\mathcal{P}$ as bitopological property (resp. p-topological property).

Utilizing Theorem 1, Corollary 1 and Corollary 2 we easily obtain the following corollary.

Corollary 5. *An i-Lindelöf property is i-topological property, a Lindelöf property and a B-Lindelöf property are both bitopological property and p-topological property.*

One question may be asked concerning the idea of n-topological space. What happens if we consider $n > 2$? This n-topological space can be constructed with the same idea, and the question in the 2-topological setup do extend naturally. A more abstract and may be less relevant generalization would be a N-topological space when X has continuum elements and N means a countable number of topologies residing on X. A difference may be arise for the last consideration.

Acknowledgement. *The authors are very grateful to the referee for his significant observations which improved the value of this paper very much.*

References

[1] B. P. Dvalishvili, Bitopological Spaces: Theory, relations with generalized algebraic structures, and applications, North-Holland Math. Stud. 199, Elsevier, 2005.
[2] Ali A. Fora and Hasan Z. Hdeib, On pairwise Lindelöf spaces, Rev. Colombiana Mat., **17**(2) (1983), 37-57.
[3] J. C. Kelly, Bitopological spaces, *Proc. London Math. Soc.*, **13**(3)(1963), 71-89.
[4] A. Kılıçman and Z. Salleh, On pairwise Lindelöf bitopological spaces, Topology & Its Appl., **154** (8) (2007), 1600-1607.
[5] R. Kopperman, The Stone-Čech compactification of a partially ordered set via bitopology, note, January 3, 2003.
[6] A. Tallafha, A. Al-Bsoul, A. Fora, Countable dense homogeneous bitoplogical spaces, *Tr. J. Math.*, 23 (1999), 233-242 © TÜBİTAK.
[7] S. Willard, General Topology, *Addison-Wesley*, Canada, 1970.

Department of Mathematics, University Malaysia Terengganu, 21030 Kuala Terengganu, Terengganu, Malaysia.
E-mail address: akilicman@umt.edu.my

Institute For Mathematical Research, University Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia.
E-mail address: bidisalleh@yahoo.com

ALBANIAN JOURNAL
OF MATHEMATICS
Volume 1, Number 3, Pages 125–131
ISSN 1930-1235: (electronic version)

PSEUDOPRIMES IN CERTAIN LINEAR RECURRENCES

FLORIAN LUCA AND IGOR E. SHPARLINSKI

(Communicated by T. Shaska)

Abstract. Let $b > 1$ be a fixed positive integer. We study the distribution of pseudoprimes to base b in certain linear recurrence sequences. We prove, in effective form, that most terms of these sequences are not pseudoprimes to base b.

1. Introduction

1.1. **Motivation.** Let $b \geqslant 2$ be an integer. Recall that a *pseudoprime to base* b is a composite positive integer m such that the congruence $b^m \equiv b \pmod m$ holds. The question of the distribution of pseudoprimes in certain sequences of positive integers has received some interest lately. For example, van der Poorten and Rotkiewicz show that any arithmetic progression $a \mod d$ with a and d coprime contains infinitely many pseudoprimes to base b; see [9] for details. Pseudoprime to base b values of the Fibonacci numbers, polynomials and the Euler function have been studied in [7], while pseudoprime Cullen and Woodall numbers are analyzed in [8]. In a recent paper, the authors jointly with Cojocaru, fixed an elliptic curve $\mathbf{E}$ defined over $\mathbb{Q}$ and studied the primes p such that the reductions of $\mathbf{E}$ modulo p are base b pseudoprimes (see [2]).

Note that Fibonacci Cullen and Woodall numbers as well as polynomials, are all examples of linearly recurrence sequences. In this paper, we continue this program and look at the presence of pseudoprimes in linear recurrence sequences of certain general types. One application of our results is an upper bound on the number of pseudoprimes amongst the numbers of $\mathbb{F}_{q^n}$-rational points on a given elliptic curve over a finite field $\mathbb{F}_q$ of q elements for $n \leqslant x$.

1.2. **The set up.** Let $u = (u_n)_{n\geqslant 0}$ be a linear recurrence sequence of integers satisfying a homogeneous linear recurrence relation

$$u_{n+k} = a_1 u_{n+k-1} + \cdots + a_{k-1} u_{n+1} + a_k u_n\,, \qquad n = 1, 2, \ldots,$$

with the characteristic polynomial

$$\psi(X) = X^k - a_1 X^{k-1} - \cdots - a_{k-1} X - a_k \in \mathbb{Z}[X].$$

Received by the editors June 13, 2007, and in revised form, July 14, 2007.
2000 *Mathematics Subject Classification.* 11N25, 11N37, 11A07.
Key words and phrases. pseudoprimes, linear recurrences.

We assume, without loss of generality, that $a_k \neq 0$. It is then well-known that

$$u_n = \sum_{i=1}^{m} A_i(n)\alpha_i^n,$$

where $\alpha_1, \ldots, \alpha_m$ are the distinct roots of $\psi(X)$, of multiplicities $\sigma_1, \ldots, \sigma_m$, respectively, and $A_i(X)$ are polynomials of degrees $\sigma_i - 1$ for $i = 1, \ldots, m$, with coefficients in $\mathbb{K} = \mathbb{Q}[\alpha_1, \ldots, \alpha_m]$.

We recall that $\alpha_1, \ldots, \alpha_m$ are also called the *characteristic roots*. Further, assume that $(u_n)_{n\geqslant 0}$ is nondegenerate, namely that α_i/α_j is not a root of 1 for any $1 \leqslant i < j \leqslant m$. It is well-known that there exist only finitely many n such that $u_n = 0$ (see, for example, [10] for a bound on the number of such n). From now on, we may assume that $n > n_0$ is large so that $u_n \neq 0$.

We refer to [4] for these and other known facts about linear recurrence sequences.

In this paper, we study the number $N_{b,u}(x)$ of positive integers $n \leqslant x$ such that u_n is a base b pseudoprime where the sequence $u = (u_n)_{n\geqslant 0}$ satisfies one additional condition.

1.3. **Divisibility sequences.** Throughout the paper, we always assume that the sequence $(u_n)_{n\geqslant 0}$ is a *divisibility sequence*, that is, $u_m \mid u_n$ whenever $m \mid n$.

By the main result in [1] (see also [3] for a more general result), we know that $u_n \mid w_n$, where $(w_n)_{n\geqslant 0}$ is a recurrence whose general term has the shape

$$w_n = an^h \prod_{j=1}^{s} \frac{\beta_j^n - \gamma_j^n}{\beta_j - \gamma_j} \tag{1}$$

for some constants $a \in \mathbb{K}$, integer $h \geqslant 0$, and algebraic integers β_j, γ_j for $j = 1, \ldots, s$ such that β_j/γ_j is not a root of unity for any $j = 1, \ldots, s$. An immediate consequence of this representation is that

$$u_n = n^{h_0} v_n, \tag{2}$$

where $(v_n)_{n\geqslant 0}$ is a linear recurrence sequence having only simple roots and $h_0 \geqslant 0$ is some integer. It is also clear that the sequence $(v_n)_{n\geqslant 0}$ is of order at most k.

Note also that $h_0 = 0$ if and only if the characteristic polynomial $\Psi(X)$ of $(u_n)_{n\geqslant 0}$ has no multiple roots.

1.4. **Examples.** Let $(F_n)_{n\geqslant 0}$ be the Fibonacci sequence given by $F_0 = 0,\ F_1 = 1$ and $F_{n+2} = F_{n+1} + F_n$ for all $n \geqslant 0$. It is well-known that the sequence $(F_n)_{n\geqslant 0}$ is a divisibility sequence. In fact, $F_n = w_n$, where $(w_n)_{n\geqslant 0}$ is given by formula (1) with $s = 1$ and β_1, γ_1 are the golden section and its conjugate. In particular, it follows from our general results, that the set of n such that F_n is a base b pseudoprime is of asymptotic density zero. As we have mentioned, this is already proved in [7]. The same remarks apply to the Pell sequence $(P_n)_{n\geqslant 0}$ given by $P_1 = 0,\ P_1 = 1$ and $P_{n+2} = 2P_{n+1} + P_n$ for all $n \geqslant 0$. Our results show that even the product F_nP_n is a base b pseudoprime only for a set of n of asymptotic density zero.

The Cullen and Woodall numbers, denoted C_n and D_n, respectively, are given by $C_n = n2^n + 1$ and $D_n = n2^n - 1$ for all $n \geqslant 1$. The sequences $(C_n)_{n\geqslant 0}$ and $(D_n)_{n\geqslant 0}$ are ternary recurrent of common characteristic polynomial $\psi(X) = (X-1)^2(X-2)$. However, none of them is a divisibility sequence so our general result does not apply to this sequence. However, using different arguments, it has been shown in [8] that

the set of n such that C_n or D_n is a base b pseudoprime is of asymptotic density zero.

The sequence of values of the Euler functions $(\varphi(n))_{n\geqslant 1}$ is a divisibility sequence because $\varphi(m) \mid \varphi(n)$ for all $m \mid n$, but it is not linearly recurrent. Nevertheless, it is shown in [7] that the set of n such that $\varphi(n)$ is a base b pseudoprime is of asymptotic density zero.

Let q be a prime power and let $\mathbf{E}$ be an ordinary elliptic curve defined over a finite field of q elements $\mathbb{F}_q$. Let $m(n)$ be the number of points on $\mathbf{E}$ defined over $\mathbb{F}_{q^n}$. Then both sequences $(m(n))_{n\geqslant 1}$ and $(m(n)/m(1))_{n\geqslant 1}$ are divisibility sequences. Indeed, $m(n) = (\tau^n - 1)(\overline{\tau}^n - 1)$, where τ and $\overline{\tau}$ are the two eigenvalues of the Frobenius. In the non-supersingular case, we know that $\tau/\overline{\tau}$ is not a root of 1 (see, for example, [6, Lemma 5]), therefore our results show that each one of the numbers $m(n)$ and $m(n)/m(1)$ is a base b pseudoprime only for a set of n of asymptotic density zero. This compliments the results of [2], where it is shown that for a fixed elliptic curve $\mathbf{E}$ over $\mathbb{Q}$, under some natural assumptions, the set of primes p such that the reductions of $\mathbf{E}$ modulo p are base b pseudoprimes forms a subset of primes of relative density zero (in the set of all primes).

1.5. **Notation.** Throughout this paper, for any positive real number x and any integer $\ell \geqslant 1$, we write $\log_\ell x$ for the function defined inductively by $\log_1 x = \max\{\ln x, 1\}$, where $\ln x$ is the natural logarithm of x, and $\log_\ell x = \log_1(\log_{\ell-1} x)$ for $\ell > 1$. When $\ell = 1$, we omit the subscript in order to simplify the notation; however, we continue to assume that $\log x \geqslant 1$ for any $x > 0$.

We use the Landau symbol O and the Vinogradov symbols $\ll$ and $\gg$ with their usual meanings, with the understanding that any implied constants depend on our data such as the sequence $(u_n)_{n\geqslant 0}$ and the number b. We recall that the notations $A \ll B$, $B \gg A$ and $A = O(B)$ are all equivalent to the fact that there exists a constant c such that the inequality $|A| \leqslant cB$ holds for all sufficiently large values of the input.

We always use the letters p and q to denote prime numbers, while m and n always denote positive integers.

1.6. **Congruences with linear recurrence sequences.** We make use of the following bound from [11] (see also [4, Theorem 5.11]).

Lemma 1. *Let $m \geqslant 2$ be an integer coprime to infinitely many elements of a nondegenerate linear recurrence sequence $(w_n)_{n\geqslant 0}$ of order k. Then for any integer $N \geqslant 1$ the number $R(N, m)$ of solutions of the congruence*

$$u(n) \equiv 0 \pmod{m}, \qquad 0 \leqslant n \leqslant N - 1,$$

satisfies the bound

$$R(N, m) \leqslant C(k)(N/\log m + 1),$$

where $C(k)$ depends only on k.

2. Main Results

2.1. **Characteristic polynomial with multiple roots.** Here we consider the case when $h_0 > 0$ in the representation (2).

Theorem 2. *Assume that a nondegenerate linear recurrence sequence $(u_n)_{n\geqslant 0}$ is a divisibility sequence. If $h_0 > 0$ in the representation* (2)*, then*

$$N_{b,u}(x) \ll \frac{x \log_2 x \log_3 x}{\log x}.$$

Proof. We let x be large and set

$$w = (\log x)^2, \qquad y = x^{1/(k+2)} \qquad \text{and} \qquad z = \exp\left(24 \log_2 x \log_3 x\right). \tag{3}$$

For a prime number p coprime to b we let $t(p)$ denote the multiplicative order of b modulo a prime p. We let $\mathcal{Q}$ be the set of primes $p \in [z, y]$ with $t(p) \leqslant p^{1/3}$. It is shown in the proof of Theorem 1 in [8] that

$$\sum_{p\in\mathcal{Q}} \frac{1}{p} \ll \frac{1}{z^{1/3}} \ll 1. \tag{4}$$

For an integer m we write $P(m)$ for the largest prime divisor of m with the convention that $P(0) = P(\pm 1) = 1$. For a prime p we put $q_p = P(t(p))$.

We define $\mathcal{R}$ as the set of primes $p \in [y, z]\backslash\mathcal{Q}$ with $q_p \leqslant w$. Clearly, each prime $p \in \mathcal{R}$ has the property that $p-1$ has a divisor $d \geqslant y^{1/3}$ with $P(d) \leqslant w$. Therefore,

$$\sum_{p\in\mathcal{R}} \frac{1}{p} \ll x \sum_{\substack{y^{1/3}\leqslant d\leqslant z\\ P(d)<w}} \sum_{\substack{p<d^3\\ p\equiv 1 \pmod d}} \frac{1}{p}.$$

The arguments used in the proof of Theorem 2 in [8] lead easily to the bound

$$\sum_{p\in\mathcal{R}} \frac{1}{p} \ll \frac{(\log_2 x)^2}{z^{(\log w)/6}} = 1. \tag{5}$$

We let $\mathcal{P}$ be the set of primes $p \in [z, y]\backslash\left(\mathcal{Q} \cup \mathcal{R}\right)$. We let $\mathcal{E}$ be the set of positive integers $n \leqslant x$ which do not have a divisor $p \in \mathcal{P}$.

By the Brun sieve inequality (see Theorem 2.2 in [5]), we have

$$\#\mathcal{E} \ll x \prod_{p\in\mathcal{P}} \left(1 - \frac{1}{p}\right).$$

Using (4) and (5), we obtain

$$\#\mathcal{E} \ll x \prod_{p\in[z,y]} \left(1 - \frac{1}{p}\right). \tag{6}$$

By Mertens's formula (see [12] for a better error term) for a positive real number t we have

$$\prod_{p\leqslant t} \left(1 - \frac{1}{p}\right) = e^{\gamma} \log t \left(1 + O\left(\frac{1}{\log t}\right)\right).$$

Applying this with $t = y$ and $t = z$ and dividing the two relations obtained in this way we get, by estimate (6),

$$\#\mathcal{E} \ll x \frac{\log z}{\log y}. \tag{7}$$

We now let $\mathcal{N}$ be the set of positive integers $n \leqslant x$ which are not in $\mathcal{E}$. Each positive integer $n \in \mathcal{N}$ has a prime factor $p \in \mathcal{P}$. To get an upper bound on $\#\mathcal{N}$,

it suffices, for every $p \in \mathcal{P}$, to count the number $M_{b,u}(x,p)$ of $m \leqslant x/p$ such that $n = mp \in \mathcal{N}$ and the congruence

$$b^{u_n} - b \equiv 0 \pmod{u_n}$$

holds. Since $h_0 > 0$, we see that

$$p \mid n \mid u_n \mid b(b^{u_n-1} - 1).$$

Clearly, if x is large enough, then

$$\gcd(q_p, a_k b D) = 1 \tag{8}$$

for all $p \in \mathcal{P}$, where we recall that $a_k = \psi(0)$ is the constant term of the characteristic polynomial $\psi(X)$ of the sequence $(u_n)_{n\geqslant 0}$, and D is the product of the discriminants of the irreducible factors of ψ. Since $p \mid b^{u_n-1} - 1$, we get

$$q_p \mid t(p) \mid u_n - 1 = n^{h_0} v_n - 1.$$

Since $q_p \mid p - 1$ and $n = mp$, we derive that

$$m^{h_0} v_{pm} \equiv 1 \pmod{q_p}. \tag{9}$$

The classical theory of linear recurrence sequences (see [4]) implies that under the condition (8) the sequence $(v_{pm})_{m\geqslant 1}$ whose order is at most k is purely periodic modulo q_p with some period $T_p \leqslant q_p^k - 1$. Furthermore, we also have the divisibility

$$T_p \mid \prod_{\nu=1}^{k} \left(q_p^\nu - 1\right),$$

which in turn implies that

$$\gcd(T_p, q_p) = 1. \tag{10}$$

Thus, if we write $m = r + T_p s$ with some integers r and s such that $0 \leqslant r < T_p$ and $0 \leqslant s \leqslant x/pT_p$, then (9) implies that

$$(r + T_p s)^{h_0} v_{pr} \equiv 1 \pmod{q_p}.$$

Thanks to the condition (10), we see that the last congruence tells us that s belongs to at most h_0 arithmetic progressions modulo q_p. Namely, this is a polynomial congruence for s modulo q_p of degree exactly h_0 since p does not divide its leading term $T_p^{h_0} v_{pr}$. Thus, s may take at most $h_0(x/(pT_p q_p) + 1)$ possible values in the interval $[0, x/pT_p]$, leading to the bound

$$M_{b,u}(x,p) \leqslant T_p h_0 \left(\frac{x}{pT_p q_p} + 1\right). \tag{11}$$

Notice that due to our choice of parameters,

$$pT_p q_p < p q_p^{k+1} < p^{k+2} \leqslant y^{k+2} = x.$$

Hence, the bound (11) simplifies to

$$M_{b,u}(x,p) \ll \frac{x}{pq_p}.$$

Using (7), we obtain

$$\begin{aligned} N_{b,u}(x) &\leqslant \#\mathcal{E} + \sum_{p\in\mathcal{P}} M_{b,u}(x,p) \ll x\frac{\log z}{\log y} + x\sum_{p\in\mathcal{P}}\frac{1}{pq_p} \\ &\ll x\frac{\log z}{\log y} + \frac{x}{w}\sum_{p\in\mathcal{P}}\frac{1}{p} \ll x\left(\frac{\log z}{\log y} + \frac{\log_2 y}{w}\right). \end{aligned}$$

Recalling the choice of w, y and z from (3), we obtain the desired result. □

2.2. **Characteristic polynomial without multiple roots.** Here, we consider the case when $h_0 = 0$ in the representation (2) and get a slightly weaker result than in Theorem 2.

Theorem 3. *Assume that a nodegenerate linear recurrence sequence $(u_n)_{n\geqslant 0}$ is a divisibility sequence. If $h_0 = 0$ in the representation (2), then*

$$N_{b,u}(x) \ll x\frac{\log_3 x}{\sqrt{\log x}}.$$

Proof. We let again x be large and we now set

$$w = \exp\left(\sqrt{\log_2 x}\right), \qquad y = x^{1/(k+2)}, \qquad z = \exp\left(12\sqrt{\log_2 x}\log_3 x\right). \tag{12}$$

We redefine the sets of primes $\mathcal{P}, \mathcal{Q}, \mathcal{R}$ as well as the sets of integers $\mathcal{E}, \mathcal{N}$ as in the proof of Theorem 2 but with the current choice of parameters w, y and z. Since we still have that

$$z \to \infty \qquad \text{and} \qquad z^{-1/6\log w}(\log_2 x)^2 = 1,$$

as $x \to \infty$, the bound (7) holds for our new choice of parameters as well.

To estimate $M_{b,u}(x,p)$, we note that congruence (9) now becomes

$$u_{pm} \equiv 1 \pmod{q_p} \tag{13}$$

Note also that that $u_n = v_n$, for $n = 1, 2, \ldots$, because $h_0 = 0$.

We now apply the bound of Lemma 1 to estimate the number of solutions of the congruence (13) with the linear recurrence sequence $(u_{np} - 1)_{n\geqslant 0}$, whose roots are now $\alpha_1^p, \ldots, \alpha_k^p$ and 1. We note that if $\alpha_1^p, \ldots, \alpha_k^p$ are not roots of unity, then Lemma 1 applies directly. Otherwise, if one of them is a root of unity ρ then all its conjugates must also be among $\alpha_1^p, \ldots, \alpha_k^p$. However, since there are no roots of unity among α_i/α_j for $1 \leqslant i < j \leqslant m$, then $\rho = \pm 1$. Thus, $u_{np} = w_n + A\rho^n$ where $(w_n)_{n\geqslant 0}$ is a linear recurrence sequence which has no roots of unity among it characteristic roots and their ratios. Thus, $(u_{2np} - 1)_{n\geqslant 0}$ and $(u_{(2n+1)p} - 1)_{n\geqslant 0}$ are nondegenerate linear recurrence sequences to which Lemma 1 applies.

Therefore,

$$M_{b,u}(x,p) \ll \frac{x}{p\log q_p} + 1 \ll \frac{x}{p\log q_p}.$$

Once again, recalling estimate (7), we obtain

$$\begin{aligned} N_{b,u}(x) &\leqslant \#\mathcal{E} + \sum_{p\in\mathcal{P}} M_{b,u}(x,p) \ll x\frac{\log z}{\log y} + x\sum_{p\in\mathcal{P}}\frac{1}{p\log q_p} \\ &\ll x\frac{\log z}{\log y} + \frac{x}{\log w}\sum_{p\in\mathcal{P}}\frac{1}{p} \ll x\left(\frac{\log z}{\log y} + \frac{\log_2 y}{\log w}\right). \end{aligned}$$

Recalling our choice of w, y and z from (12), we obtain the desired result. □

Acknowledgement

We thank the referee for suggestions which improved the quality of the manuscript. Research of F. L. was supported by grant SEP-CONACyT 46755 that of I. S. by ARC grant DP0556431.

References

[1] J.-P. Bézivin, A. Pethö and A. J. van der Poorten, 'A full characterisation of divisibility sequences', *Amer. J. Math.* **112** (1990), no. 6, 985–1001.
[2] A. C. Cojocaru, F. Luca and I. E. Shparlinski, 'Pseudoprime reductions of elliptic curves', *Preprint*, 2007.
[3] P. Corvaja and U. Zannier, 'Finiteness of integral values for the ratio of two linear recurrences', *Invent. Math.* **149** (2002), no. 2, 431–451.
[4] G. Everest, A. van der Poorten, I. E. Shparlinski and T. Ward, *Recurrence sequences*, Mathematical Surveys and Monographs, **104**, American Mathematical Society, Providence, RI, 2003.
[5] H. Halberstam and H.-E. Richert, *Sieve methods*, Academic Press, London, 1974.
[6] F. Luca and I. E. Shparlinski, 'On the exponent of the group of points on elliptic curves in extension fields', *Int. Math. Res. Not.* **2005**, no. 23, 1391–1409.
[7] F. Luca and I. E. Shparlinski, 'Pseudoprime values of the Fibonacci sequence, polynomials and the Euler function', *Indag. Math.* **17** (2006), 611–625.
[8] F. Luca and I. E. Shparlinski, 'Pseudoprime Cullen and Woodall numbers', *Colloq. Math.* **107** (2007), 35–43.
[9] A. J. van der Poorten and A. Rotkiewicz, 'On strong pseudoprimes in arithmetic progressions', *J. Austral. Math. Soc. Ser. A* **29** (1980), no. 3, 316–321.
[10] H. P. Schlickewei and W. M. Schmidt, 'The number of solutions of polynomial-exponential equations', *Compositio Math.* **120** (2000), no. 2, 193–225.
[11] I. E. Shparlinski, 'The number of different prime divisors of recurrence sequences', *Matem. Zametki* **42** (1987), 494–507 (in Russian).
[12] A. I. Vinogradov, 'On the remainder in Mertens's formula,' *Dokl. Akad. Nauk SSSR* **148** (1963), 262–263 (in Russian).

Florian Luca, Instituto de Matemáticas, Universidad Nacional Autónoma de México, C.P. 58089, Morelia, Michoacán, México
E-mail address: fluca@matmor.unam.mx

Igor E. Shparlinski, Dept. of Computing, Macquarie University, Sydney, NSW 2109, Australia
E-mail address: igor@ics.mq.edu.au

ALBANIAN JOURNAL
OF MATHEMATICS
Volume 1, Number 3, Pages 133–143
ISSN 1930-1235: (electronic version)

EXTENDING TOPOLOGICAL GROUP ACTIONS TO CONFORMAL GROUP ACTIONS

A. WOOTTON

Abstract. A consequence of the resolution of the Nielsen Realization Problem is that if G is any finite topological group of automorphisms of a compact oriented surface S, then there exists a complex structure on S so that the action of G extends to a conformal action on S. We show that this complex structure is unique if and only if G is a triangle group, and for all other groups, there are infinitely many such actions.

1. Introduction

Suppose that G is a group of homeomorphisms of a compact, connected, oriented surface S of genus $g \geqslant 2$. A consequence of the resolution of the Nielsen Realization Problem (NRP), see [8], is that there exists a group of homeomorphisms of S which is topologically equivalent to G (which by abuse of notation we call G) and some complex structure on S so that the action of G extends to a conformal action on S. It is natural to ask for a given finite group of homeomorphisms G of S, how many such complex stuctures exist which extend the action of G to a conformal action, and are there any groups for which the struture is unique? In most circumstances, the complex structure is not unique, and in fact there are infinitely many different complex structures which can be imposed on S so that G acts conformally (see Lemma 3.3). However, in the special case that G is a triangle group, there are only finitely many conformal structures which can be imposed on S so that G acts conformally (see Lemma 3.2). We shall strengthen this result and show that if G is a triangle group, then there is a unique structure which can be imposed on S so that G extends to a conformal action (see Theorem 3.6). A consequence of our results is an enumeration method for the number of different topological group actions of a triangle group G on a surface S in terms of surface kernel epimorphisms (for which we may utilize techniques developed such as those in [6] or [13]).

Following the NRP, there has been tremendous progress in the study of topological equivalence classes of group actions on surfaces due to the interactions with conformal group actions, see for example [3], [11], [12]. For details on this, see Section 2, or [3] for a more thorough exposition. One of the motivational reasons for this is that there is a one-one correspondence between finite subgroups of the mapping class group $\mathcal{M}_\sigma$ of a surface of genus σ and the topological equivalence classes of finite groups of homeomorphisms which can act on a surface of genus σ. In general, the problem of enumerating conjugacy classes of subgroups of $\mathcal{M}_\sigma$ for arbitrary σ is highly computational and depends very much upon how a group

Received by the editors June 30, 2007 and in revised form, August 30, 2007.
2000 *Mathematics Subject Classification.* Primary: 14H45, 14H37, 14H30, 20F34.
Key words and phrases. Mapping Class Group, Quasiplatonic Surface, Triangle Group.

G acts on a surface S as well as the general structure of G. For triangle groups however, we have much more control over how G may act on S, and so a general enumeration formula is much more realistic (see Proposition 4.4). Another important fact about triangle groups is that many such groups will be maximal as finite subgroups of $\mathcal{M}_\sigma$, and for those which are not, there are computational methods to determine precisely which ones are not maximal, see for example [5] or [9]. Thus an enumeration method for the number of different topological group actions which are triangle groups can be used to provide a lower bound on the number of maximal finite subgroups of $\mathcal{M}_\sigma$, thus providing insight into the general structure of $\mathcal{M}_\sigma$.

Our paper is structured as follows. In Secion 2 we develop the necessary preliminary results regarding topological group actions summarizing the results from [3]. We shall also outline the preliminaries for counting the conformal equivalence classes of complex structures which can be imposed on a surface of genus g. Following this, we shall prove the main result in Section 3. Though on initial consideration, it may seem to be a highly computational result, the proof is suprisingly straight forward with the combination of some classical results and more recent techniques. In Section 4, we use some more recent results to determine an enumeration method to count the number of such groups. We finish by presenting some explicit examples. We note that throughout the paper, by surface we mean a compact, oriented, 2-manifold.

2. Preliminaries on Topological and Conformal Group Actions

Let G be a finite group. The group G is said to act topologically (in an orientation preserving manner) on surface a S of genus $\sigma \geqslant 2$ if there is an injection

$$\varepsilon : G \hookrightarrow \mathrm{Homeo}^+(S)$$

into the group of orientation preserving homeomorphisms (we shall identify G with its image under ε). Two actions $\varepsilon_1, \varepsilon_2$ are said to be *topologically equivalent* if there is a homeomorphism h of S and an automorphism ω of G such that

$$\varepsilon_2(\omega(g)) = h \circ \varepsilon_1(g) \circ h^{-1}.$$

This is equivalent to saying that the images $\varepsilon_1(G)$ and $\varepsilon_2(G)$ are conjugate in $\mathrm{Homeo}^+(S)$.

For $\sigma \geqslant 2$, due to the NRP, Fuchsian groups provide us with a way to examine topological group actions. Specifically, a surface S of genus $\sigma \geqslant 2$ is topologically equivalent to a quotient of the upper half plane $\mathbb{H}/\Lambda$ where Λ is any torsion free Fuchsian group isomorphic to the fundamental group of S called a surface group for S. A finite group G acts on S if and only if G is topologically equivalent to Γ/Λ for some Fuchsian group Γ containing such a Λ as a normal subgroup of index $|G|$. The structure of Γ is completely determined by the ramification data of the quotient map $\pi_G \colon S \to S/G$ which must satisfy the Riemann-Hurwitz formula. Specifically, if the quotient map π_G branches over r points with ramification indices m_i for $1 \leqslant i \leqslant r$ and the quotient space S/G has genus g, then a presentation for Γ is:

$$\Gamma = \langle a_1, b_1, \ldots, a_g, b_g, c_1, \ldots, c_r | c_1^{m_1}, \ldots, c_r^{m_r}, \prod_{i=1}^{r} c_i \prod_{j=1}^{g} [a_j, b_j] \rangle \tag{1}$$

where

$$\sigma = 1 + |G|(g-1) + \frac{|G|}{2}\sum_{i=1}^{r}\left(1 - \frac{1}{m_i}\right).$$

Such a group is described by the tuple $(g; m_1, \ldots, m_r)$ called the **signature** of Γ (we also say that G has signature $(g; m_1, \ldots, m_r)$). In the special case that $g = 0$ and $r = 3$, we call G and Γ **triangle groups**. Such group actions are usually described through the use of surface kernel epimorphisms, so we interpret our observations accordingly.

Theorem 2.1. *A finite group G acts on a surface S of genus $\sigma \geqslant 2$ with signature $(g; m_1, \ldots, m_r)$ if and only if there exists a Fuchsian group with signature $(g; m_1, \ldots, m_r)$ and an epimorphism $\varrho\colon \Gamma \to G$ preserving the orders of the elements of finite order (called a surface kernel epimorphism) such that*

$$\sigma = 1 + |G|(g-1) + \frac{|G|}{2}\sum_{i=1}^{r}\left(1 - \frac{1}{m_i}\right).$$

A useful way to describe surface kernel epimorphisms is through the use of generating vectors defined as follows (see [3]).

Definition 2.2. A vector of group elements $(\alpha_1, \beta_1, \ldots, \alpha_g, \beta_g, \eta_1, \ldots, \eta_r)$ in a finite group G is called a $(g; m_1, \ldots, m_r)$-generating vector for G if all of the following hold:

(*i*) $G = \langle \alpha_1, \beta_1, \ldots, \alpha_g, \beta_g, \eta_1, \ldots, \eta_r \rangle$
(*ii*) $\Pi_{i=1}^{g}[\alpha_i, \beta_i] \cdot \Pi_{j=1}^{r}\eta_j = 1$ (where $[,]$ denotes the commutator).
(*iii*) $O(c_i) = m_i$ (where $O(.)$ denotes group order).

Clearly any $(g; m_1, \ldots, m_r)$-generating vector $\mathcal{V}$ for G defines a unique surface kernel epimorphism from a fixed Γ with signature $(g; m_1, \ldots, m_r)$ onto G, called the surface kernel epimorphism of $\mathcal{V}$. Conversely, any surface kernel epimorphism $\varrho\colon \Gamma \to G$ uniquely defines a generating vector called the generating vector of ϱ. Thus topological group actions can be described through the utilization of generating vectors of finite groups. The exact correspondence is given in the following.

Theorem 2.3. *Two equivalence classes of $(g; m_1, \ldots, m_r)$-generating vectors of the finite group G define the same topological equivalence class of G-actions if and only if the generating vectors lie in the same $\mathrm{Aut}\,(G) \times \mathrm{Aut}\,(\Gamma)$-class where the action of $\mathrm{Aut}\,(G) \times \mathrm{Aut}\,(\Gamma)$ on a generating vector $\mathcal{V}$ is defined by the action on a surface kernel epimorphism ϱ with generating vector $\mathcal{V}$ given by $(\alpha, \gamma) \cdot \varrho = \alpha \circ \varrho \circ \gamma^{-1}$ for $\alpha \in \mathrm{Aut}\,(G)$, $\gamma \in \mathrm{Aut}\,(\Gamma)$.*

Proof. See Proposition 2.2 of [3]. □

Thus given a generating vector $\mathcal{V}$ for a group G, it defines a topological action, namely the action of Γ/Λ on $S = \mathbb{H}/\Lambda$ where Λ is the kernel of the surface kernel epimorphism of $\mathcal{V}$ where Γ is a Fuchsian group with signature $(g; m_1, \ldots, m_r)$ (note that Γ can be any subgroup of $\mathrm{PSL}\,(2, \mathbb{R})$ with signature $(g; m_1, \ldots, m_r)$). Conversely, given a group acting topologically on S with signature $(g; m_1, \ldots, m_r)$, it defines a $(g; m_1, \ldots, m_r)$-generating vector up to $\mathrm{Aut}\,(G) \times \mathrm{Aut}\,(\Gamma)$ equivalence. Specifically, it defines the class containing the generating vector of $\varrho\colon \Gamma \to G$ where ϱ is the surface kernel epimorphism from Γ with signature $(g; m_1, \ldots, m_r)$ and kernel Λ such that S is topologically equivalent to $\mathbb{H}/\Lambda$ and G is topologically

equivalent to Γ/Λ. We call any generating vector from this $\mathrm{Aut}(G) \times \mathrm{Aut}(\Gamma)$ class **a generating vector of** G.

Now observe that given a surface kernel epimorphism $\varrho\colon \Gamma \to G$ with kernel Λ, there exists a complex structure which can be imposed on $S = \mathbb{H}/\Lambda$ so that the group G acts conformally on S - namely the natural structure inherited from the complex structure on $\mathbb{H}$. This motivates the following definition.

Definition 2.4. Suppose G is a finite group of homeomorphisms of a surface S and G is topologically equivalent to Γ/Λ where Γ and Λ are Fuchsian groups with Λ isomorphic to the fundamental group of S so S is topologicially equivalent to $\mathbb{H}/\Lambda$. Then we say the complex structure which can be imposed on S inherited from the complex structure on $\mathbb{H}$ **extends the action of** G **to a conformal action on** S.

Unlike toplogical actions, this structure depends upon a choice for Γ. Specifically, we have the following.

Theorem 2.5. *Two surfaces $S_1 = \mathbb{H}/\Lambda_1$ and $S_2 = \mathbb{H}/\Lambda_2$ are conformally equivalent if and only if Λ_1 and Λ_2 are conjugate in* $\mathrm{PSL}(2,\mathbb{R})$.

In particular, given a topological group action G on S, there may exist multiple inequivalent structures on S so that the action of G extends to a conformal action on S. Indeed, our observations imply the following.

Theorem 2.6. *Suppose G is a group acting on a surface S of genus $\sigma \geqslant 2$ with generating vector $\mathcal{V}$. Then the number of conformal structures which can be imposed on S so that the action of G extends to a conformal action is equal to the number of* $\mathrm{PSL}(2,\mathbb{R})$*-conjugacy classes of kernels of surface kernel epimorphism whose generating vector is $\mathcal{V}$.*

Thus in order to determine the number of structures which can be imposed on S so that G with generating vector $\mathcal{V}$ extends to a conformal action on S, we need to count $\mathrm{PSL}(2,\mathbb{R})$-classes of surface subgroups which are normal in all Fuchsian groups with signature given by the generating vector $\mathcal{V}$. It is a well known fact that the set of conjugacy classes of groups with signature $(g; m_1, \ldots, m_r)$ is homeomorphic to $\mathbb{R}^{6g-6+2r}$ (see [1]). This fact coupled with Theorem 2.6 suggests that for a given a group G with generating vector $\mathcal{V}$ acting on S, unless G is a triangle group, there will be an infinite number of structures which can be imposed on S so that G extends to a conformal action, and in the special case that G is a triangle group, there are finitely many such structures. This leads to the following questions which are the central focus of our work .

Question 2.7. Suppose G is a group acting on a surface S of genus $\sigma \geqslant 2$ with generating vector $\mathcal{V}$. How many conformal structures do there exist which can be imposed on S so that G extends to a conformal action on S? In addition, are there any groups actions for which there exists a unique conformal structure extending the action of G to a conformal action?

3. The Main Result

We shall prove our main result through a series of Lemmas. Henceforth, assume that G is a group acting on S with generating vector $\mathcal{V}$ and, where relevant, that G is topologically Γ/Λ for Fuchsian groups Γ and Λ where Λ is isomorphic to the fundamental group of S. We need the following important result regarding

$\mathrm{PSL}(2,\mathbb{R})$-conjugacy classes of Fuchsian groups and the existence of overgroups (see [1] and [10]).

Theorem 3.1. *Suppose that Γ is a Fuchsian group with signature $\mathcal{S}$.*

- (*i*) *If G is a triangle group, there is a unique $\mathrm{PSL}(2,\mathbb{R})$-conjugacy class of Fuchsian groups with signature $\mathcal{S}$.*
- (*ii*) *If G is not a triangle group, there exist infinitely many $\mathrm{PSL}(2,\mathbb{R})$-conjugacy classes of Fuchsian groups with signature $\mathcal{S}$. Moreover, we have the following additional information about the existence of overgroups:*
  - (a) *If the signature of Γ does not appear in Singermans list,* [10]*, there are infinitely many $\mathrm{PSL}(2,\mathbb{R})$-conjugacy classes of finitely maximal Fuchsian groups with signature $\mathcal{S}$ (Γ is finitely maximal if there is no Fuchsian group Δ with $\Gamma \leqslant \Delta$ and $[\Delta : \Gamma] < \infty$).*
  - (b) *If the signature of Γ does appear in Singermans list,* [10]*, there is a list of signatures $\mathcal{L}$ such that given any group Γ with signature $\mathcal{S}$ and any signature $\mathcal{S}_\Delta \in \mathcal{L}$, there is a Fuchsian group Δ with signature $\mathcal{S}_\Delta$ with $\Gamma \leqslant \Delta$ and $[\Delta : \Gamma] < \infty$. Moreover, there exist infinitely many $\mathrm{PSL}(2,\mathbb{R})$-conjugacy classes of subgroups with signature $\mathcal{S}$ such that if Γ is any such given group with signature $\mathcal{S}$, the only possible signatures for a group Δ with $\Gamma \leqslant \Delta$ and $[\Delta : \Gamma] < \infty$ are those from $\mathcal{L}$.*

We can use this result to show that the only possible candidates are triangle groups.

Lemma 3.2. *If G is a triangle group with generating vector $\mathcal{V}$ acting on S, then there are finitely many structures which can be imposed on S extending the action of G to a conformal action.*

Proof. Suppose that G is a triangle group. We shall show that there exists only finitely many structures extending the action of G to a conformal action independent of the generating vector for G and hence there can only be finitely many once a generating vector has been specified.

If G has signature $(0; m_1, m_2, m_3)$, let Γ be a fixed Fuchsian group with signature $(0; m_1, m_2, m_3)$. We first observe that since all triangle groups are conjugate in $\mathrm{PSL}(2,\mathbb{R})$, given any surface kernel Λ_1 such that $\Gamma_1/\Lambda \cong G$ where Γ_1 also has signature $(0; m_1, m_2, m_3)$, there exists a surface kernel $\Lambda \leqslant \Gamma$ with $\Gamma/\Lambda \cong G$ which is $\mathrm{PSL}(2,\mathbb{R})$-conjugate to Λ_1. It follows that the number of $\mathrm{PSL}(2,\mathbb{R})$-conjugacy classes of surface subgroups of Fuchsian groups with signature $(0; m_1, m_2, m_3)$ and quotient group G is bounded above by the number of surface kernels of the fixed Fuchsian group Γ with quotient group G. Since G is finite, there are only finitely many surface kernel epimorphisms from Γ to G, so only finitely many surface kernels with quotient group G and hence only finitely many conformal structures which can be imposed on S so that G acts conformally.

□

Lemma 3.3. *If a group G with generating vector $\mathcal{V}$ acting on S is not a triangle group, then there exist infinitely many different structures which can be imposed on S so that the action of G extends to a conformal action on S.*

Proof. Suppose that G with signature $(g; m_1, \ldots, m_r)$ and generating vector $\mathcal{V}$ acting on S is not a triangle group. If the signature of G does not appear in Singermans

list, Theorem 3.1 implies there exist infinitely many $\mathrm{PSL}(2,\mathbb{R})$-conjugacy classes of finitely maximal subgroups of $\mathrm{PSL}(2,\mathbb{R})$ with signature $(g; m_1, \ldots, m_r)$. For such a Γ, let $\varrho\colon \Gamma \to G$ denote a surface kernel epimorphism with generating vector $\mathcal{V}$. Since automorphism groups of compact Riemann surfaces of genus $g \geqslant 2$ are finite, it follows that Γ is the normalizer of $\mathrm{Ker}(\varrho)$ in $\mathrm{PSL}(2,\mathbb{R})$. It follows that if two finitely maximal groups Γ_1 and Γ_2 with signature $(g; m_1, \ldots, m_r)$ are not $\mathrm{PSL}(2,\mathbb{R})$-conjugate, and $\varrho_1\colon \Gamma_1 \to G$ and $\varrho_2\colon \Gamma_2 \to G$ are surface kernel epimorphisms with generating vector $\mathcal{V}$, then the kernels $\mathrm{Ker}(\varrho_1)$ and $\mathrm{Ker}(\varrho_2)$ are not $\mathrm{PSL}(2,\mathbb{R})$-conjugate (since normalizers of conjugate subgroups of a group are conjugate). Since there are infinitely many different $\mathrm{PSL}(2,\mathbb{R})$-conjugacy classes of finitely maximal subgroups of $\mathrm{PSL}(2,\mathbb{R})$ with signature $(g; m_1, \ldots, m_r)$, Theorem 2.6 implies there are infinitely many different structures which can be imposed on G such that the action of G extends to a conformal action.

Now suppose that the signature of G appears in Singermans list and let $\mathcal{L}$ be the list of signatures such that given any Γ with signature $(g; m_1, \ldots, m_r)$, for each signature $\mathcal{S} \in \mathcal{L}$, there always exists an overgroup Δ of Γ with signature $\mathcal{S}$ and $[\Delta : \Gamma] < \infty$. Now for a fixed epimorphism $\varrho\colon \Gamma \to G$, Theorem 5.1 of [5] gives complete conditions for $\mathrm{Ker}(\varrho)$ to be normal in $\Delta \geqslant \Gamma$ with signature from $\mathcal{L}$. In particular, these conditions are dependent only on ϱ and either hold true for all possible Γ with signature $(g; m_1, \ldots, m_r)$ or none. Let $\mathcal{S}_\Delta$ denote the signature of the largest Fuchsian group with signature from $\mathcal{L}$ in which any $\mathrm{Ker}(\varrho)$ is normal (this is well defined by our remarks). Observe that since signature determines a group up to isomorphism, if Δ is a group with signature $\mathcal{S}_\Delta$, there always exists a subgroup Γ with signature $(g; m_1, \ldots, m_r)$ and a normal surface subgroup Λ with $\Lambda \lhd \Gamma$ such that $\eta\colon \Gamma \to \Gamma/\Lambda \cong G$ is $\mathrm{Aut}(G)$ (and hence $\mathrm{Aut}(G) \times \mathrm{Aut}(\Gamma)$) equivalent to ϱ as specified above. To finish, we note that since any group with signature from $\mathcal{L}$ is not a triangle group (since Γ is not a triangle group), Theorem 3.1 implies there exist infinitely many $\mathrm{PSL}(2,\mathbb{R})$-classes of subgroups with signature $\mathcal{S}_\Delta$ which can only be subgroups of Fuchsian groups with signature from $\mathcal{L}$. In particular, if Δ is a group with signature $\mathcal{S}_\Delta$ from one of these classes and Γ and Λ are as specified above, then $\Delta = N(\Lambda)$, so we can apply an identical argument to the previous case and the result follows.

□

This result means that the only possible candidates for which there exists a unique structure extending the action of G are triangle groups. Since we are looking for $\mathrm{Aut}(G) \times \mathrm{Aut}(\Gamma)$-classes of generating vectors, we need to examine the group $\mathrm{Aut}(\Gamma)$ in more detail when Γ is a triangle group.

Theorem 3.4. *Suppose that* Γ *has signature* $(0; m_1, m_2, m_3)$.

(*i*) *If all the* m_i *are distinct, then* $\mathrm{Aut}(\Gamma) = \mathrm{Inn}(\Gamma)$.
(*ii*) *If precisely two of the* m_i *are equal, then* $[\mathrm{Aut}(\Gamma) : \mathrm{Inn}(\Gamma)] = 2$.
(*iii*) *If all of the* m_i *are equal, then* $[\mathrm{Aut}(\Gamma) : \mathrm{Inn}(\Gamma)] = 6$.

Proof. A consequence of [4] is that if F is the free group on two generators, then $[\mathrm{Aut}(F) : \mathrm{Inn}(F)] = 6$. Since the automorphisms of Γ will be the same as the automorphisms of F which preserve orders of elements, the result follows. Alternatively, we could construct the $\mathrm{Aut}(\Gamma)$ explicitly using the generators of general mapping class groups given in [2]. □

Lemma 3.5. *Suppose* Γ *is a Fuchsian triangle group. Then there exists another Fuchsian group* Δ *containing* Γ *as a normal subgroup and an isomorphism* $\Phi\colon \Delta \to \mathrm{Aut}\,(\Gamma)$ *induced by the action of conjugation of* Δ *on* Γ.

Proof. Suppose that Γ has signature $(0; m_1, m_2, m_3)$ and let Δ be a Fuchsian group with $\Gamma \lhd \Delta$. We shall first show that the map $\Phi\colon \Delta \to \mathrm{Aut}\,(\Gamma)$ induced by the action of conjugation of the elements of Δ on Γ is injective. To see this, suppose that conjugation by $\gamma \in \Delta$ and $\delta \in \Delta$ induce the same automorphism. Then it follows that for all $c \in \Gamma$, $\gamma c \gamma^{-1} = \delta c \delta^{-1}$, or equivalently $\delta^{-1}\gamma$ commutes with every element in Γ. However, two non-identity elements in a Fuchsian group Γ commute if and only if they have the same fixed point set (see for example Theorem 5.2.4 of [7]). If $\delta^{-1}\gamma$ is non-trivial, it follows that all the elements of Γ have the same fixed point set and hence Γ is commutative which is not true. Hence $\delta^{-1}\gamma$ is trivial so $\delta = \gamma$, so the map Φ is injective.

To finish the proof, we observe that under the map Φ induced by conjugation, $\Phi(\Gamma) = \mathrm{Inn}(\Gamma)$ (the inner automorphism group of Γ). The result then follows through observation of the different possible overgroups of Γ given in Singermans list, [10]. Specifically if all the m_i are distinct, then $\mathrm{Aut}\,(\Gamma) = \mathrm{Inn}(\Gamma)$ so the result trivially holds. If precisely two of the m_i equal, there exists a Fuchsian group Δ with $\Gamma \lhd \Delta$ and $[\Delta : \Gamma] = 2$, and if all the m_i's are equal, then there exists a Fuchsian group Δ with $\Gamma \lhd \Delta$ and $[\Delta : \Gamma] = 6$.

□

We are now ready to prove our main result.

Theorem 3.6. *There exists a unique conformal structure extending the action of a topological group of automorphisms* G *to a conformal action if and only if* G *is a triangle group. For all other groups, there exist infinitely many different structurse extending the action of* G *to a conformal action.*

Proof. If G is not a triangle group, Lemma 3.3 proves the result. Therefore, suppose G is a triangle group with signature $\mathcal{S}$ acting on a surface S of genus $\sigma \geqslant 2$ with generating vector $\mathcal{V}$. By Theorem 2.6, we need to show that there is just one $\mathrm{PSL}\,(2, \mathbb{R})$-conjugacy class of kernels of surface kernel epimorphism whose generating vector $\mathrm{Aut}\,(G) \times \mathrm{Aut}\,(\Gamma)$-equivalent to $\mathcal{V}$. To prove this, we shall first show that for a fixed triangle group Γ with signature $\mathcal{S}$, all kernels of surface kernel epimorphisms with generating vector $\mathrm{Aut}\,(G) \times \mathrm{Aut}\,(\Gamma)$-equivalent to $\mathcal{V}$ are $\mathrm{PSL}\,(2, \mathbb{R})$-conjugate. We shall then show that if Γ_1 is any other triangle group with signature $\mathcal{S}$, any kernel of a surface kernel epimorphism from Γ_1 to G with generating vector $\mathrm{Aut}\,(G) \times \mathrm{Aut}\,(\Gamma)$-equivalent to $\mathcal{V}$ is $\mathrm{PSL}\,(2, \mathbb{R})$-conjugate to one of the surface kernels in Γ.

Suppose that Γ is some fixed triangle group with signature $\mathcal{S}$ and let $\mathcal{K}$ denote the set of all kernels of surface kernel epimorphisms from Γ to G with generating vector $\mathrm{Aut}\,(G) \times \mathrm{Aut}\,(\Gamma)$-equivalent to $\mathcal{V}$. If $\Lambda_1, \Lambda_2 \in \mathcal{K}$, let $\varrho_1, \varrho_2\colon \Gamma \to G$ denote corresponding surface kernel epimorphisms. Since the generating vectors of ϱ_1, ϱ_2 are $\mathrm{Aut}\,(G) \times \mathrm{Aut}\,(\Gamma)$-equivalent to $\mathcal{V}$, they are $\mathrm{Aut}\,(G) \times \mathrm{Aut}\,(\Gamma)$-equivalent to each other and so $\varrho_1 = \alpha \circ \varrho_2 \circ \gamma^{-1}$ for some $\alpha \in \mathrm{Aut}\,(G)$ and $\gamma \in \mathrm{Aut}\,(\Gamma)$. It follows that $\mathrm{Ker}\,(\varrho_1) = \mathrm{Ker}\,(\varrho_2 \circ \gamma^{-1})$, or equivalently $\gamma(\Lambda_2) = \mathrm{Ker}\,(\Lambda_1)$. However, by Lemma 3.5, every element of $\mathrm{Aut}\,(\Gamma)$ is induced by the action of conjugation by some overgroup Δ of Γ in $\mathrm{PSL}\,(2, \mathbb{R})$, and in particular, it follows that Λ_1 and Λ_2 are $\mathrm{PSL}\,(2, \mathbb{R})$-conjugate. Thus for a fixed triangle group Γ, any two surface

kernel epimorphisms with $\mathrm{Aut}\,(G) \times \mathrm{Aut}\,(\Gamma)$-equivalent generating vectors have $\mathrm{PSL}\,(2,\mathbb{R})$-conjugate kernels.

Now suppose $\Gamma_1 \neq \Gamma$ is a triangle group with signature $\mathcal{S}$ and suppose that $\eta\colon \Gamma_1 \to G$ is a surface kernel epimorphism with generating vector $\mathrm{Aut}\,(G) \times \mathrm{Aut}\,(\Gamma)$-equivalent to $\mathcal{V}$. To finish the proof, we need to show that $\mathrm{Ker}\,(\eta)$ is $\mathrm{PSL}\,(2,\mathbb{R})$-conjugate to a group in the set $\mathcal{K}$. Since Γ_1 and Γ are triangle groups, there exists $T \in \mathrm{PSL}\,(2,\mathbb{R})$ such that $T\Gamma_1 T^{-1} = \Gamma$. Now T induces a conformal map between the quotient spaces (see for example Theorem 5.9.3 of [7])

$$\bar{T}\colon \mathbb{H}/\,\mathrm{Ker}\,(\eta) \to \mathbb{H}/(T\,\mathrm{Ker}\,(\eta)T^{-1})$$

and with this map, we have $\bar{T}(\Gamma_1/\,\mathrm{Ker}\,(\eta))\bar{T}^{-1} = \Gamma/(T\,\mathrm{Ker}\,(\eta)T^{-1})$. Since $\bar{T}$ is conformal, it is a homeomorphism, and so it follows that the groups $\Gamma_1/\,\mathrm{Ker}\,(\eta)$ and $\Gamma/(T\,\mathrm{Ker}\,(\eta)T^{-1})$ are topologically equivalent. In particular, the generating vector of the map $\varrho\colon \Gamma \to G$ with kernel $T\,\mathrm{Ker}\,(\eta)T^{-1}$ will be $\mathrm{Aut}\,(G) \times \mathrm{Aut}\,(\Gamma)$-equivalent to $\mathcal{V}$ and so $T\,\mathrm{Ker}\,(\eta)T^{-1} \in \mathcal{K}$, hence the result.

□

4. Enumeration of Topological Equivalence Classes of Triangle Group Actions

Our results in the previous section imply the following result relating the number of topological equivalence classes of group actions of triangle groups and the $\mathrm{PSL}\,(2,\mathbb{R})$-conjugacy classes of subgroups of triangle groups.

Proposition 4.1. *The number of topological equivalence classes of group actions of a group G with signature $(0; m_1, m_2, m_3)$ on S of genus $\sigma \geqslant 2$ is equal to the number of $N(\Gamma)$-conjugacy classes of surface subgroups of a fixed Fuchsian triangle group Γ with quotient group G.*

Thus in order to determine the number of topological equivalence classes of group actions of a triangle group G with signature $(0; m_1, m_2, m_3)$, we just need to count the number of surface kernels in a Fuchsian group Γ with signature $(0; m_1, m_2, m_3)$ up to conjugation in the normalizer of Γ in $\mathrm{PSL}\,(2,\mathbb{R})$. The next result provides a way to determine $N(\Lambda) \cap N(\Gamma)$ dependent on the maps defined below.

Definition 4.2. Suppose (x, y, z) is a generating vector for a triangle group G. Then we define the following identifications:

- $i_1\colon x \mapsto y, y \mapsto x, z \mapsto yzy^{-1}$
- $i_2\colon x \mapsto y^{-1}xy, y \mapsto z, z \mapsto y$
- $i_3\colon x \mapsto z, y \mapsto xyx^{-1}, z \mapsto x$
- $j\colon x \mapsto y, y \mapsto z, z \mapsto x$

Theorem 4.3. *Suppose that Λ is a surface kernel subgroup of a triangle group Γ and $\Gamma \lhd \Delta$. Let $\varrho\colon \Gamma \to G = \Gamma/\Lambda$ be the corresponding surface kernel epimorphism and suppose (x, y, z) is the generating vector of ϱ.*

(i) If Γ has signature $(0; m_1, m_1, m_2)$, then Λ is normal in Δ with signature $(0; 2, m_1, 2m_2)$ if and only if the identification i_1 extends to an automorphism of G.

(ii) If Γ has signature $(0; m, m, m)$, then Λ is normal in a group Δ with signature $(0; 2, m, 2m)$ and is normal in no larger group if and only if only precisely one of the identifications i_1, i_2, or i_3 extends to an automorphism of G

(*iii*) *If* Γ *has signature* $(0; m, m, m)$*, then* Λ *is normal in a group* Δ *with signature* $(0; 3, 3, m)$ *and is normal in no larger group if and only if the identification* j *extends to an automorphism of* G *but the identification* i_1 *does not.*

(*iv*) *If* Γ *has signature* $(0; m, m, m)$*, then* Λ *is normal in a group* Δ *with signature* $(0; 2, 3, 2m)$ *if and only if the identifications* i_1 *and* j *extend to automorphisms of* G*.*

Proof. This is a consequence of the results developed in [5].

□

Putting our results together, we get the following.

Proposition 4.4. *The number of topologically inequivalent topological* G*-actions* $\mathcal{T}$ *with signature* $(0; m_1, m_2, m_3)$ *on a surface* S *can be calculated as follows.*

(*i*) *If all the* m_i *are distinct,*

$$\mathcal{T} = \frac{|\mathcal{V}_G|}{|\operatorname{Aut}(G)|}$$

where $\mathcal{V}_G$ *denotes the set of all generating vectors for surface kernel epimorphisms from* Γ *to* G*.*

(*ii*) *If* Γ *has signature* $(0; m_1, m_1, m_2)$ *with* $m_2 \neq m_1$*, then*

$$\mathcal{T} = \frac{|\mathcal{V}_G|}{2|\operatorname{Aut}(G)} + \frac{|\mathcal{V}_{G,i_1}|}{|\operatorname{Aut}(G)|}$$

where $\mathcal{V}_G$ *denotes the set of generating vectors of surface kernel epimorphisms from* Γ *to* G *for which the identification* i_1 *does not extend to an automorphism of* G *and* $\mathcal{V}_{G,i_1}$ *denotes the set of generating vectors for which the map* i_1 *does extend to an automorphism of* G*.*

(*iii*) *If* Γ *has signature* $(0; m, m, m)$*, then the number of topological equivalence classes of group actions of* G *on a surface* S *with signature* $(0; m, m, m)$ *is equal to*

$$\mathcal{T} = \frac{|\mathcal{V}_G|}{6|\operatorname{Aut}(G)|} + \frac{|\mathcal{V}_{G,i}|}{3|\operatorname{Aut}(G)|} + \frac{|\mathcal{V}_{G,j}|}{2|\operatorname{Aut}(G)|} + \frac{|\mathcal{V}_{G,i,j}|}{|\operatorname{Aut}(G)|}$$

where $\mathcal{V}_G$ *denotes the set of generating vectors of surface kernel epimorphisms from* Γ *to* G *for which none of the transformations* i_1, i_2, i_3 *or* j *extend to automorphisms of* G*,* $\mathcal{V}_{G,i}$ *denotes the set of generating vectors for which just one of* i_1, i_2 *or* i_3 *extends to an automorphism of* G*,* $\mathcal{V}_{G,j}$ *denotes the set of generating vectors for which* j *extends to an automorphism, but* i_1 *does not, and* $\mathcal{V}_{G,i,j}$ *denotes the set of generating vectors for which* i_1 *and* j *extend to automorphisms of* G*.*

Proof. For a given Γ, the proof is a simple enumeration of the size of the orbits of surface kernels under the action of conjugation by subgroups of $N(\Gamma)$.

□

We finish with some examples.

Example 4.5. Suppose G is cyclic of order 7 generated by x and Γ has signature $(0; 7, 7, 7)$. Then any surface kernel epimorphism from Γ with signature $(0; 7, 7, 7)$ has corresponding generating vector of the form (x^a, x^b, x^{7-a-b}), so there are 30 epimorphisms in total. For each of these generating vectors, we need to determine

which identifications given in Definition 4.2 extend to automorphisms of G. Observe that if (x, y, z) is any generating vector with a repeated entry, then switching the two repeated entries induces an automorphism of G (namely the identity), so either i_1, i_2 or i_3 extends to an automorphism of G. Now note that since there can be at most repeated entry, if (x, y, z) is a generating vector with a repeated entry, then the identification j does not extend to an automorphism of G. Since there are 18 such generating vectors, we have $|\mathcal{V}_{G,i}| = 18$.

If a generating vector (x, y, z) has three distinct entries, we can use a similar argument to show that the identification j always extends to an automorphism of G but the identification i_1 does not. Since there are 12 such generating vectors, we have $|\mathcal{V}, |_G| = 18$. There are no other generating vectors, so applying Proposition 4.4, we get

$$\mathcal{T} = \frac{|\mathcal{V}_{G,i_1}|}{3|\operatorname{Aut}(G)|} + \frac{|\mathcal{V}_{G,j}|}{2|\operatorname{Aut}(G)|} = \frac{18}{18} + \frac{12}{12} = 2.$$

Therefore there are two distinct topological group actions of the cyclic group of order 7 with signature $(0; 7, 7, 7)$ on a surface of genus 3.

As we briefly mentioned in the introduction, there are numerous computational methods developed to help calculate the size of the sets of generating vectors (see for example [6] or [13]). We illustrate with an example.

Example 4.6. Suppose that $G = C_{13} \rtimes C_3$ (the non-trivial semidirect product). Then G acts on a surface S of genus 6 with signature $(0; 3, 3, 13)$. Applying Theorem 3 of [6], we get at most 156 different generating vectors for G, but since the smallest group containing an element of order 3 and order 13 is 39, it follows that all these generating vectors are generating vectors for G. Since $|\operatorname{Aut}(G)| = 78$, it follows that there are 2 different surface kernels in Γ with signature $(0; 3, 3, 13)$ with quotient G. Since there are no groups which act on a surface S of genus 6 with signature $(0; 2, 3, 26)$, it follows that these two surface kernels are conjugate by an element in the Fuchsian $\Delta \geqslant \Gamma$ with signature $(0; 2, 3, 26)$. Hence there is just one $\mathrm{PSL}(2, \mathbb{R})$-class of surface kernels and thus a unique topological action of G on S with signature $(0; 3, 3, 13)$.

References

[1] L. Ahlfors *On quasiconformal mappings.* J. Analyse Math. 3, (1954). 1–58.

[2] J. Birman. *Braids, links, and mapping class groups.* Annals of Mathematics Studies, No. 82. Princeton University Press, 1974.

[3] S.A. Broughton, *Classifying Finite Group Actions on Surfaces of Low Genus*, J. Pure and Appl. Alg., Vol. 69 (1990) pp. 233–270.

[4] B. Chang. *The Automorphism Group of the Free Group with Two Generators.* Michigan Mathematics Journal, 7, Iss. 1, (1960), pp. 79-81

[5] E. Bujalance, F. J. Cirre, M. D. E. Conder. *On Extendability of Group Actions on Compact Riemann Surfaces.* Trans. Amer. Math. Soc. 355 (2003), 1537-1557.

[6] G. A. Jones. *Enumeration of Homomorphisms and Surface Coverings.* Quarterly J. Math. Oxford (2) 46 (1995), 485-507.

[7] G. Jones, D. Singerman. *Complex functions. An algebraic and geometric viewpoint.* Cambridge University Press, Cambridge, 1987.

[8] S. Kerckhoff, *The Nielsen Realization Problem.* Annals of Math., Vol. 117 (1983), pp. 235–265.

[9] K. Magaard, T. Shaska, S. Shpectorov, H. Voelklein, *The locus of curves with prescribed automorphism group.* RIMS Kyoto Series, Communications on Arithmetic Fundamental Groups, vol. 6, 112–141, 2002.

[10] David Singerman. *Finitely Maximal Fuchsian Groups.* J. London Math. Soc., (2), 6(1972),29-38.
[11] E. Tyszkowska. *Topological classification of conformal actions on elliptic-hyperelliptic Riemann surfaces.* J. Algebra 288 (2005), no. 2, 345–363
[12] E. Tyszkowska. *Topological classification of conformal actions on 2-hyperelliptic Riemann surfaces.* Bull. Inst. Math. Acad. Sinica 33 (2005), no. 4, 345–368.
[13] Aaron Wootton. *Counting Belyĭ p-gonal Surfaces with Many Automorphisms.* Proc. 10th International Conf. on Appl. of Comp. Alg., (2004).

University of Portland, 5000 North Willamette Blvd., Portland, OR 97203
E-mail address: wootton@up.edu

ALBANIAN JOURNAL
OF MATHEMATICS
Volume 1, Number 3, Pages 145–155
ISSN 1930-1235: (electronic version)

POLYNOMIAL COMPLEXITY FOR HILBERT SERIES OF BOREL TYPE IDEALS

AMIR HASHEMI

Abstract. In this paper, it is shown that the Hilbert series of a Borel type ideal may be computed within a complexity which is polynomial in D^n where $n+1$ is the number of unknowns and D is the highest degree of a minimal generator of input (monomial) ideal.

1. Introduction

A classical algorithm to compute the Hilbert series of a monomial ideal, is from its free resolution which is infeasible in practice. Bayer and Stillman [2] have proved that the computation of Hilbert series of a monomial ideal is at least difficult as an NP-complete problem in the number of variables, see also [5]. For some class of monomial ideals, the computation of the Hilbert series may be less costly than NP-complete. For example, Bayer and Stillman [2] have shown that the Hilbert series of a Borel monomial ideal may be computed in linear time. Recall that a monomial ideal J over the ring $R = K[x_0, \ldots, x_n]$ where K is an arbitrary field is defined to be Borel if $x_j m \in J$ implies that $x_i m \in J$ for any $i < j$.

In this paper, we study the complexity of computing the Hilbert series of a *Borel type ideal*. A monomial ideal $J \subset R$ is Borel type if it satisfies the following property:

$$J : x_j^{\infty} = J : \langle x_0, \ldots, x_j \rangle^{\infty}$$

for all $j = 1, \ldots, n$ (see [1, 12]). *We show that the Hilbert series of a Borel type ideal may be computed within a complexity which is polynomial in D^n where $n+1$ is the number of unknowns and D is the highest degree of a minimal generator of input polynomials.* For this, we describe an algorithm to decide within the same complexity whether a monomial ideal is Borel type or not. Also, we establish a sharper upper bound for the satiety and Castelnuovo-Mumford regularity of such an ideal and we prove that these invariants may be computed within the above complexity. Finally, as an application of our results, we give a new formula to compute the degree of a Borel type ideal. This paper is a continuation of the ideas which have first appeared in [11].

It is well-known that to compute the Hilbert series of a general ideal, we reduce it to a monomial ideal by Gröbner basis computation. On the other hand, it follows from the work of Mayr and Meyer [14] that the problem of computing a Gröbner basis (in worst case) is exponential space complete. Both cardinality and maximal

Received by the editors June 8, 2007 and, in revised form, August 31, 2007.
2000 *Mathematics Subject Classification.* Primary: 13D45, Secondary: 13P10, 13F20.
Key words and phrases. Ideals, Borel ideals, Hilbert series .

degree of a Gröbner basis might be doubly exponential in the number of variables. Our result shows that (for an ideal whose initial ideal is Borel type) if the problem of computing a Gröbner basis is simple then that of Hilbert series is not more difficult. In fact, our computation suggests that the expensive part of computing the Hilbert series of a general ideal is the Gröbner basis computation. This leads us to the following conjecture which generalizes our result.

Conjecture 1.1. *The Hilbert series of a monomial ideal may be computed within a complexity which is polynomial in D^n where $n+1$ is the number of unknowns and D is the highest degree of a minimal generator of the input ideal.*

The main interest of our algorithm is its bound of complexity. In fact, with the existing implementations (see [5, 2]), the computation of the Hilbert series is negligible with respect to that of the Gröbner basis which is needed for. It is therefore not worthwhile to spent human time to efficiently implement our algorithm.

Now, we give the structure of the paper. In Sections 2, we recall the definition of a Borel type ideal and we describe a polynomial-time algorithm to test whether a monomial ideal is Borel type or not. In Section 3 (resp. 4) we establish an upper bound and describe an algorithm having the complexity polynomial in D^n for the satiety (resp. Castelnuovo-Mumford regularity) of a Borel type ideal. In Section 5, we prove that one can compute the Hilbert series of a Borel type ideal within this complexity. In Section 6, we give a formula to compute the degree of a Borel type ideal. Finally, Section 7 presents our conclusions.

2. Borel type ideals

The purpose of this section is to study a certain class of monomial ideal which we call *Borel type ideals*. We describe also an algorithm which determines whether a monomial ideal is Borel type or not within a complexity which is polynomial in input size.

We recall first the definition of the saturation of a homogeneous ideal. Let J be a monomial ideal of the polynomial ring $R = K[x_0, \ldots, x_n]$ where K is an arbitrary field. If $\mathfrak{m} = \langle x_0, \ldots, x_n \rangle$ is the unique maximal homogeneous ideal of R, then we recall that the ideal $J : \mathfrak{m}^i$ is defined for any positive integer i as

$$J : \mathfrak{m}^i = \{F \in R \mid \forall G \in \mathfrak{m}^i,\ GF \in J\}.$$

The ideal $J : \mathfrak{m}^\infty$ is defined $\bigcup_{i=1}^{\infty} J : \mathfrak{m}^i$. Denote by J_ℓ the set of homogeneous elements of degree ℓ of J.

Proposition 2.1. *The ideal $J^{\text{sat}} = J : \mathfrak{m}^\infty$ is called the* saturation *of J. It is the unique largest ideal $I \subset R$ for the following property:*

$$\exists s \text{ such that } \forall \ell \geq s \quad I_\ell = J_\ell.$$

Proof. One can check easily that the saturation of J satisfies this property. □

For a monomial ideal J, we introduce the following sequences of ideals associated to J. Let $R_i = K[x_0, \ldots, x_i]$.

Notation 2.2. *Let* $\text{sec}(J, 0) = \overline{\text{sec}}(J, 0) = J$ *and for* $i = 1, \ldots, n+1$*:*

- $\text{sec}(J, i) = J + \langle x_{n-i+1}, \ldots, x_n \rangle$.
- $\overline{\text{sec}}(J, i) = \text{sec}(J, i) \cap R_{n-i} = J|_{x_{n-i+1} = \cdots = x_n = 0} \cap R_{n-i}$.

Note that $\text{sec}(J, i)$ and $\overline{\text{sec}}(J, i)$ are ideals of R and R_{n-i} respectively for any i.

Lemma 2.3. *Let $J \subset R$ be a monomial ideal. For any $\ell \leq n$, the following conditions are equivalent:*

(1) $\mathrm{sec}(J,i)^{\mathrm{sat}} = \mathrm{sec}(J,i) : x_{n-i}^{\infty}$ *for* $i = 0, \ldots, \ell$.
(2) $\overline{\mathrm{sec}}(J,i)^{\mathrm{sat}} = \overline{\mathrm{sec}}(J,i) : x_{n-i}^{\infty}$ *for* $i = 0, \ldots, \ell$.
(3) $J : \langle x_0, \ldots, x_{n-i}\rangle^{\infty} = J : x_{n-i}^{\infty}$ *for* $i = 0, \ldots, \ell$.

Proof. (1) $\Rightarrow$ (3). We proceed by induction on i. For $i = 0$, we have $I^{\mathrm{sat}} = I : x_n^{\infty}$ by definition of sec, and this proves the assertion in this case. Suppose that the assertion is true for $i-1$. We have to prove that any (monomial) minimal generator $m \in J : x_{n-i}^{\infty}$ belongs to $J : \langle x_0, \ldots, x_{n-i}\rangle^{\infty}$. For some k, we have $x_{n-i}^k m \in J \subset \mathrm{sec}(J,i)$. Thus, $m \in \mathrm{sec}(J,i)^{\mathrm{sat}}$ by (1), and therefore $x_j^t m \in \mathrm{sec}(J,i) = J + \langle x_{n-i+1}, \ldots, x_n\rangle$ for some integer t and for any j. We claim that $m \notin \langle x_{n-i+1}, \ldots, x_n\rangle$. If this claim is true, $x_j^t m \in J$ for $j = 0, \ldots, n-i$, and this proves the assertion.

Proof of the claim: We prove it ad absurdum. Let $m = x_j m'$ for some $j \in \{n-i+1, \ldots, n\}$. Thus $m' \in J : x_j^{\infty} = J : \langle x_0, \ldots, x_j\rangle^{\infty}$ by the hypothesis of induction. This implies that $m' \in J : x_{n-i}^{\infty}$. Since m is a minimal generator of $J : x_{n-i}^{\infty}$, this is impossible.

(3) $\Rightarrow$ (1). It is enough to prove that any monomial $m \in \mathrm{sec}(J,i) : x_{n-i}^{\infty}$ belongs to $\mathrm{sec}(J,i)^{\mathrm{sat}}$ for any i. Two cases are possible: If m belongs to $\langle x_{n-i+1}, \ldots, x_n\rangle$ then there is nothing to prove because $\langle x_{n-i+1}, \ldots, x_n\rangle \subset \mathrm{sec}(J,i)^{\mathrm{sat}}$. If not, there exists an integer k such that $x_{n-i}^k m \in J$. This implies that $m \in J : x_{n-i}^{\infty} = J : \langle x_0, \ldots, x_{n-i}\rangle^{\infty}$ by (3). Thus, there exists an integer t such that $x_j^t m \in J$ for $j = 0, \ldots, n-i$, and therefore $x_j^t m \in \mathrm{sec}(J,i)$ for any j. This implies that $m \in \mathrm{sec}(J,i)^{\mathrm{sat}}$. This argument was inspired by the proof of [4], Proposition 3.2.

(2) $\Rightarrow$ (3). The proof is similar to (1) $\Rightarrow$ (3).

(3) $\Rightarrow$ (2). It suffices to show that any monomial $m \in \overline{\mathrm{sec}}(J,i) : x_{n-i}^{\infty}$ belongs to $\overline{\mathrm{sec}}(J,i)^{\mathrm{sat}}$ for any i. We have $x_{n-i}^k m \in J$ for some integer k. Thus, $m \in J : x_{n-i}^{\infty} = J : \langle x_0, \ldots, x_{n-i}\rangle^{\infty}$ by (2) which implies that there exists an integer t such that $x_j^t m \in J$ for $j = 0, \ldots, n-i$. The membership $x_j^t m \in R_{n-i}$ proves the assertion. □

We recall that the dimension $\dim(J)$ of the ideal J is the dimension of the corresponding quotient ring.

Lemma 2.4. *If any condition of Lemma 2.3 is true for $\ell = \dim(J) - 1$, it is true for any ℓ.*

Proof. By Lemma 2.3, it is enough to prove the assertion for the first condition. Notice that if X is a zero-dimensional (monomial) ideal then $X^{\mathrm{sat}} = X : x_i^{\infty}$ for any i. Now apply this for $X = \mathrm{sec}(J,i)$ which is zero-dimensional by definition. □

Definition 2.5. *A monomial ideal $J \subset R$ is called a* Borel type ideal *if it satisfies one of the equivalent conditions in Lemma 2.3 for $\ell = \dim(J) - 1$.*

From Lemma 2.3(3), we conclude that the notions Borel type and nested type (introduced in [4]) coincide.

Lemma 2.3(2) provides a new characterization of Borel type ideals from which we derive a simple test for determining whether a monomial ideal is Borel type or not (see the following).

Algorithm testing Borel type ideal

Input: $J \subset R$ a monomial ideal
Output: The answer to "Is J a Borel type ideal?"
$G := \{m_1, \ldots, m_k\}$ a minimal system of generators for J
$Deg := \max\{\deg(m_1), \ldots, \deg(m_k)\}$
$e :=$ highest integer ℓ such that $x_i^{Deg} \in J$ for $i = 0, \ldots, \ell$
$d := n - e$
For each monomial $x_0^{e_0} \cdots x_h^{e_h} \in G$ with $h > n - d$ and $e_h > 0$ **do**
  For $j = 1, \ldots, h-1$ **do**
  If $x_0^{e_0} \cdots x_{h-1}^{e_{h-1}} x_j^{Deg} \notin J$ **then**
  Return "No"
Return "Yes, and the dimension of the ideal is d"

Proof. (Algorithm) The termination of the algorithm is obvious. Let us show its correctness. For this, we have to prove that J is Borel type if and only if the response of the algorithm is "Yes". Suppose that J is Borel type and $x_0^{e_0} \cdots x_h^{e_h} \in G$ for some $h > n - d$ with $e_h > 0$. This implies that (Lemma 2.3(2))

$$x_0^{e_0} \cdots x_{h-1}^{e_{h-1}} \in J : x_h^{\infty} = J : \langle x_0, \ldots, x_h \rangle^{\infty}.$$

Thus, $x_0^{e_0} \cdots x_{h-1}^{e_{h-1}} x_j^{Deg}$ must be in J for any $j = 0, \ldots, h-1$, and the answer is "Yes". In this case, d is the dimension of J by Nœther normalization test (see [3], Lemma 3.1). Conversely, we can conclude that $J : x_h^{\infty} \subset J : \langle x_0, \ldots, x_h \rangle^{\infty}$, and therefore J is Borel type (Lemma 2.3(2)).

□

Remark 2.6. *The condition $h > n - d$ in this algorithm is not essential, because $x_i^{Deg} \in J$ for $i = 0, \ldots, n - d$.*

Remark 2.7. *The integer d is the dimension of J if the answer of the algorithm is "Yes" (see the proof of algorithm).*

Proposition 2.8. *The complexity of this algorithm is polynomial in kn.*

Proof. One can easily see that the number of operations in two loops "For" is k^2n^2. Thus the complexity of the algorithm is polynomial in kn. □

3. Satiety of Borel type ideals

In this section, we prove a new upper bound for the satiety of a Borel type ideal, and we describe an algorithm of polynomial complexity in input size that computes the satiety of such an ideal. Let us define the satiety of a monomial ideal.

Definition 3.1. *The* satiety *of a monomial ideal $J \subset R$, denoted by* $\operatorname{sat}(J)$*, is the smallest positive integer s such that $J_\ell^{\mathrm{sat}} = J_\ell$ for all $\ell \geq s$.*

We show first an upper bound for the satiety of a Borel type ideal. For this, a lemma from [1] is needed. Here, a linear form $y \in R$ is generic for J if y is a non-zero divisor in R/J^{sat}.

Lemma 3.2. *Let $J \subset R$ be a monomial ideal and $y \in R$ be a linear form. The following conditions are equivalent:*

(1) $(J : y)_\ell = J_\ell$ *for any* $\ell \geq s$.

(2) $\mathrm{sat}(J) \leq s$ *and* y *is generic for* J.

Corollary 3.3. *Let* $J \subset R$ *be a Borel type ideal. Then*

$$\mathrm{sat}(J) = \max_{m \in (J:x_n)\setminus J} \{\deg(m)\} + 1.$$

Proof. Since x_n is generic for J from hypothesis, $(J : x_n)_\ell = J_\ell$ for any $\ell \geq \mathrm{sat}(J)$ by Lemma 3.2. Thus, the satiety of J is equal to the maximum degree of $(J : x_n)\setminus J$ plus one which proves the assertion.

□

The following theorem may follow from [4], Corollary 2.6. We give here a simpler proof for it.

Theorem 3.4. *Let* $J \subset R$ *be a Borel type ideal and let* $x_0^{D_0} \cdots x_n^{D_n}$ *be the least common multiple of the minimal generators of* J. *Then,*

$$\mathrm{sat}(J) \leq \max\{0, D_0 + \cdots + D_n - n\}.$$

Proof. Two cases are possible: If $D_0 + \cdots + D_n - n < 0$, there is some i such that $D_i = 0$. We claim that $\mathrm{sat}(J) = 0$. For this it is enough to show that $J : \mathfrak{m} = J$, i.e. J is saturated. Let $m \in J : \mathfrak{m}$ be a monomial. Thus, $x_i m \in J$ and this implies that $m \in J$ because x_i does not appear in the generators of J and this proves the claim. In the other case, by Corollary 3.3, it suffices to prove that any monomial $m \in J : x_n$ of degree $D_0 + \cdots + D_n - n$ belongs to J. From degree of m, one can show that $x_i^{D_i}$ divides m for some i. The membership $x_n m \in J$ implies that $x_i^t m \in J$ for some t because J is Borel type. This follows that $m \in J$ by the fact that $x_i^{D_i}$ divides m and D_i is the maximal degree ℓ such that x_i^ℓ appears in the minimal generators of J. □

Example 3.5. Computing the upper bounds for the satieties of some Borel type ideals. *Let* R *be the ring* $K[x_0, x_1, x_2, x_3, x_4]$. *Consider the monomial ideal* $J = \langle x_0, x_1\rangle$. *Since* $1+1-4 < 0$, *then* $\mathrm{sat}(J) = 0$. *The satiety of* $J = \langle x_0^2, x_1^4, x_2^5, x_3^3, x_4\rangle$ *is less than or equal to* 11 *because* $2 + 4 + 5 + 3 + 1 - 4 \geq 0$.

The following lemma is the basis for the occurrence of D^n in our complexity bounds.

Lemma 3.6. *The number of monomials of degree at most* $\delta = (n + 1)(D - 1) + 1$ *in* $n + 1$ *variables is bounded above by* $(eD)^{n+1}$ *for* D *and* $n \geq 0$.

Proof. The number of monomials of degree at most δ in $n + 1$ variables is equal to $\binom{n+1+\delta}{n+1}$ (see [6] p. 106 for example). By definition of the binomial coefficients, we have

$$\begin{aligned}\binom{n+1+\delta}{n+1} &= \binom{(n+1)D+1}{n+1} \\ &= \frac{D^{n+1}}{(n+1)!}\prod_{i=1}^{n+1}\left(n+1+\frac{2-i}{D}\right).\end{aligned}$$

As $n + \frac{2-i}{D} \leq n$ for $i \geq 2$ and $\left(n+1-\frac{1}{D}\right)\left(n+1+\frac{1}{D}\right) < (n+1)^2$, we have for $n > 1$

$$\binom{n+1+\delta}{n+1} < \frac{(n+1)^{n+1}}{(n+1)!}D^{n+1},$$

which implies that $\binom{n+1+\delta}{n+1} < (eD)^{n+1}$ by Stirling's formula where $e = 2.71828\cdots$ is the usual Euler constant. For $n = 0$ the result is easily proved directly. □

Corollary 3.7. *The satiety of a Borel type ideal may be computed by a complexity polynomial in D^n where D is the highest degree of its minimal generator.*

Proof. Let $J \subset R$ be a Borel type ideal. If $D = 1$, two cases are possible: If $D_n = 1$ then $J = \mathfrak{m}$ and $\mathrm{sat}(J) = 1$. In the other case, i.e. $D_n = 0$, we have $\mathrm{sat}(J) = 0$ (see Theorem 3.4). Thus, if $D = 1$ the bound polynomial in D^n holds. Now, suppose that $D \geq 2$. By Corollary 3.3 and Theorem 3.4, it is enough to find the maximal degree $h \leq \delta = (n+1)(D-1)+1$ such that there is a monomial m of this degree with $m \in (J : x_n) \setminus J$. The number of these monomials is $(eD)^{n+1}$ (Lemma 3.6) and the cost of the last condition is $k(n+1)$ operations where k is the number of minimal generators of J. Thus the complexity of this computation is $k(n+1)(eD)^{n+1}$. This is polynomial in D^n because $k \leq (eD)^{n+1}$ (by Lemma 3.6), $n+1 \leq 2^{0.55(n+1)} \leq D^{0.55(n+1)}$ and $e < D^{1.45}$ (by $D \geq 2$). □

4. Castelnuovo-Mumford regularity of Borel type ideals

In this section, we prove a new upper bound for the Castelnuovo-Mumford regularity of a Borel type ideal, and we describe an algorithm having the polynomial complexity in input size to compute the Castelnuovo-Mumford regularity of such an ideal. Let us define the *Castelnuovo-Mumford regularity* of a monomial ideal $J \subset R$. If

$$0 \longrightarrow \bigoplus_j R(e_{rj}) \longrightarrow \cdots \longrightarrow \bigoplus_j R(e_{1j}) \longrightarrow \bigoplus_j R(e_{0j}) \longrightarrow J \longrightarrow 0$$

is a minimal graded free resolution of J, $\mathrm{reg}(J)$ is the maximal of $e_{ij} - i$ for each i and j. To establish an upper bound for the Castelnuovo-Mumford regularity of a Borel type ideal, we use the following lemmas from [1].

Lemma 4.1. *Let $J \subset R$ be a monomial ideal, and $y \in R$ be a generic linear form for J. The following conditions are equivalent:*

- $\mathrm{reg}(J) \leq s$.
- $\mathrm{sat}(J) \leq s$ *and* $\mathrm{reg}(J + \langle y \rangle) \leq s$.

Lemma 4.2. *Let $J \subset R$ be a zero–dimensional monomial ideal. The following conditions are equivalent:*

- $\mathrm{sat}(J) \leq s$.
- $\mathrm{reg}(J) \leq s$.
- *J_s is equal to the set of homogeneous polynomials of degree s of R.*

Corollary 4.3. *With the same hypothesis, we have* $\mathrm{reg}(J) = \mathrm{sat}(J)$.

Proposition 4.4. *Let $J \subset R$ be a Borel type ideal and let $d = \dim(J)$.*

$$\mathrm{reg}(J) = \max_{0 \leq i \leq d} \{\mathrm{sat}(\mathrm{sec}(J, i))\} \tag{1}$$

$$= \max_{0 \leq i \leq d} \{\mathrm{sat}(\overline{\mathrm{sec}}(J, i))\}. \tag{2}$$

Proof. To prove the equality (1), from Lemma 4.1 we have

$$\mathrm{reg}(J) = \max\{\mathrm{sat}(J), \mathrm{reg}(\mathrm{sec}(J, 1))\}.$$

By reusing this formula for the ideal $\mathrm{sec}(J,1)$ and using the fact that x_{n-1} is generic for $\mathrm{sec}(J,1)$ we obtain

$$\begin{aligned}\mathrm{reg}(J) &= \max\{\mathrm{sat}(J), \max\{\mathrm{sat}(\mathrm{sec}(J,1)), \mathrm{reg}(\mathrm{sec}(J,2))\}\}\\ &= \max\{\mathrm{sat}(J), \mathrm{sat}(\mathrm{sec}(J,1)), \mathrm{reg}(\mathrm{sec}(J,2))\}.\end{aligned}$$

So by induction, we can conclude that $\mathrm{reg}(J)$ is equal to

$$\max\{\mathrm{sat}(J), \mathrm{sat}(\mathrm{sec}(J,1)), \ldots, \mathrm{sat}(\mathrm{sec}(J,d-1)), \mathrm{reg}(\mathrm{sec}(J,d))\}.$$

Since $\mathrm{sec}(J,d)$ is zero–dimensional (see for example [10], Lemma 5), then

$$\mathrm{reg}(\mathrm{sec}(J,d)) = \mathrm{sat}(\mathrm{sec}(J,d)),$$

(Corollary 4.3), and this proves the assertion.

Let us prove (2). It is enough to show that $\mathrm{sat}(\overline{\mathrm{sec}}(J,i)) = \mathrm{sat}(\mathrm{sec}(J,i))$ for any i. By the membership $x_{n-i+1}, \ldots, x_n \in \mathrm{sec}(J,i)$ we have

$$\overline{\mathrm{sec}}(J,i)^{\mathrm{sat}} = \mathrm{sec}(J,i)^{\mathrm{sat}}|_{x_{n-i+1}=\cdots=x_n=0} \cap R_{n-i},$$

by definition of the saturation of an ideal, and this proves the assertion.

□

As a consequence of Proposition 4.4 and Corollary 3.7 we have:

Corollary 4.5. *The Castelnuovo-Mumford regularity of a Borel type ideal may be computed by a complexity polynomial in D^n where D is the highest degree of its minimal generator.*

If $D = 1$, we can conclude simply that the regularity of the ideal is 1 (see the proof of Corollary 3.7). Imran and Sarfraz [13], have proved the upper bound $(n+1)D - n$ for the Castelnuovo-Mumford regularity of a Borel type ideal where D is the highest degree of a minimal generator of the ideal. We improve this bound in the following theorem.

Theorem 4.6. *Let $J \subset R$ be a Borel type ideal and let $x_0^{D_0} \cdots x_n^{D_n}$ be the least common multiple of the minimal generators of J. Then,*

$$\mathrm{reg}(J) \leq \max\{D_0 + \cdots + D_{n-d} - (n-d), \ldots, D_0 + \cdots + D_n - n\}.$$

Proof. Since $\overline{\mathrm{sec}}(J,i)$ for any i is Borel type, then its satiety is at most $\max\{0, D_0 + \cdots + D_{n-i} - n + i\}$ by Theorem 3.4. Thus, the assertion follows from Proposition 4.4 and the fact that $D_0 + \cdots + D_{n-d} > n - d$. □

Example 4.7 (Computing an upper bound for the Castelnuovo-Mumford regularity of a Borel type ideal.)**.** *Let R be the ring $K[x_0, x_1, x_2, x_3, x_4]$. Consider the monomial ideal $J = \langle x_0, x_1^2 \rangle$. Thus, its regularity is at most* $\max\{1+2-4, 1+2-3, 1+2-2, 1+2-1\} = 2$.

5. Hilbert series of Borel type ideals

In this section, we describe an algorithm to compute the Hilbert series of a Borel type ideal within a polynomial complexity in input size.

Let X be a graded module or an ideal and δ be a positive integer. We denote by X_δ (resp. $X_{\geq\delta}$) the set of elements of X of degree (resp. at least) δ. Recall that the *Hilbert series* of a monomial ideal $J \subset R$ is the power series $\mathrm{HS}_J(t) =$

$\sum_{s=0}^{\infty} \mathrm{HF}_J(s)t^s$ where $\mathrm{HF}_J(s)$, the Hilbert function of J, is the dimension of $(R/J)_s$ as a K-vector space. From this definition, we have

$$\mathrm{HS}_{J_{<\delta}}(t) = \mathrm{HF}_J(0) + \cdots + \mathrm{HF}_J(\delta-1)t^{\delta-1} + \sum_{i=\delta}^{\infty} \binom{n+i}{n} t^i \tag{3}$$

$$\mathrm{HS}_{J_{\geq\delta}}(t) = \sum_{i=0}^{\delta-1} \binom{n+i}{n} t^i + \mathrm{HF}_J(\delta)t^{\delta} + \mathrm{HF}_J(\delta+1)t^{\delta+1} + \cdots \tag{4}$$

for any δ. Thus, we can conclude that $\mathrm{HS}_J = \mathrm{HS}_{J_{<\delta}} + \mathrm{HS}_{J_{\geq\delta}} - \mathrm{HS}_{\langle 0\rangle}$. Therefore, to prove that the Hilbert series of a Borel type ideal J may be computed within a complexity polynomial in D^n, it is enough to prove the same for these three Hilbert series for $\delta = D_0 + \cdots + D_n - n$. Let us to compute $\mathrm{HS}_{J_{\geq\delta}}$. For this, we prove first that for any monomial ideal $J \subset R$, the homogeneous ideal $J_{\geq\delta}$ is *stable*. A monomial ideal $J \subset R$ is called stable (see [7]) if for all monomial $m \in J$ we have $x_j m/x_\ell \in J$ for all $j < \ell$ where ℓ is the maximal integer i such that x_i divides m. We remark that this result has proved by Imran and Sarfraz [13] for $\delta = (n+1)D-n$ and we generalize it here by another approach.

Lemma 5.1. *Let $J \subset R$ be a Borel type ideal and let $x_0^{D_0}\cdots x_n^{D_n}$ be the least common multiple of the minimal generators of J. Then $J_{\geq\delta}$ is stable for $\delta = \max\{D_0 + \cdots + D_i - i \mid i = n-d, \ldots, n\}$ with $d = \dim(J)$.*

Proof. Let $m \in J_{\geq\delta}$ be a monomial, ℓ be the maximal integer i such that x_i divides m and $j < \ell$ be an integer. By definition $m \in \overline{\mathrm{sec}}(J, n-\ell)$. It follows by the definition of Borel ideal that $m/x_\ell \in \overline{\mathrm{sec}}(J, n-\ell) : x_{n-\ell}^{\infty} \subset \overline{\mathrm{sec}}(J, n-\ell)^{\mathrm{sat}}$. This implies that $x_j m/x_\ell \in \overline{\mathrm{sec}}(J, n-\ell)$ because $\deg(x_j m/x_\ell) \geq \delta$ is greater than $\mathrm{sat}(\overline{\mathrm{sec}}(J, n-\ell))$ by Theorem 4.6 and Proposition 4.4. Therefore $x_j m/x_\ell \in J$. Remark that for $\ell < n-d$ we used the fact that the ideal $\overline{\mathrm{sec}}(J, n-\ell)$ contains $\overline{\mathrm{sec}}(J, d)$ which is zero-dimensional and thus $\mathrm{sat}(\overline{\mathrm{sec}}(J, n-\ell)) \leq \mathrm{sat}(\overline{\mathrm{sec}}(J, d)) \leq \delta$. □

Recall that $\mathrm{HS}_J(t) = P(t)/(1-t)^{n+1}$ where $P(t)$ is a polynomial in t (see [8], Theorem 7 of Chapter 11). We denote this polynomial by $\mathrm{NHS}_J(t)$.

Lemma 5.2. *With the hypotheses of Lemma 5.1, we have $\mathrm{NHS}_{J_{\geq\delta}}(t) = 1 - t^{\delta}\sum_{i=0}^{n} a_i(1-t)^i$ where a_i is the number of monomial $m \in J_\delta$ such that i is the maximal integer ℓ with $x_\ell \mid m$.*

Proof. Let $J_{\geq\delta} = \langle m_1, \ldots, m_k\rangle$ and m_i be arranged such that m_{i+1} is greater than m_i for all i using the lexicographic order $x_n > x_{n-1} > \cdots > x_0$. Thus, by [2] Corollary 2.3 we have

$$\mathrm{NHS}_{J_{\geq\delta}}(t) = \mathrm{NHS}_{\langle m_1\rangle}(t) + \sum_{i=2}^{k} t^{\deg(m_i)}\mathrm{NHS}_{\langle m_1,\ldots,m_{i-1}\rangle : m_i}(t).$$

Notice that by the stable property of $J_{\geq\delta}$ (Lemma 5.1) we have that $\langle m_1, \ldots, m_{i-1}\rangle : m_i = \langle x_0, \ldots, x_{v_i-1}\rangle$ where v_i is the maximal integer ℓ with $x_\ell \mid m_i$. Therefore, using the fact that $\deg(m_i) = \delta$

$$\mathrm{NHS}_{J_{\geq\delta}}(t) = \mathrm{NHS}_{\langle m_1\rangle}(t) + t^{\delta}\sum_{i=2}^{k} \mathrm{NHS}_{\langle x_0,\ldots,x_{v_i-1}\rangle}(t).$$

The zero-dimensionality of $J + \langle x_n, \ldots, x_{n-\dim(J)+1}\rangle$ implies that $m_1 = x_0^\delta$, and therefore $\mathrm{NHS}_{\langle m_1\rangle}(t) = 1-t^\delta$. Thus, the assertion follows from $\mathrm{NHS}_{\langle x_0,\ldots,x_{v_i-1}\rangle}(t) = (1-t)^{v_i}$ (see [2] Corollary 2.5) and definition of a_i. □

Theorem 5.3. *Let $J \subset R$ be a Borel type ideal. The Hilbert series of J may be computed by a complexity polynomial in D^n where D is the highest degree of its minimal generator.*

Proof. If $D = 1$ then $J = \langle x_0, \ldots, x_i\rangle$ for some integer i (by definition), and $\mathrm{HS}_J(t) = 1/(1-t)^{n-i}$. Thus, the bound polynomial in D^n holds in this case. Now, let $D \geq 2$ and $\delta = \max\{D_0 + \cdots + D_i - i \mid i = n - \dim(J), \ldots, n\}$ where $x_0^{D_0}\cdots x_n^{D_n}$ is the least common multiple of the minimal generators of J. Using the formula $\mathrm{HS}_J = \mathrm{HS}_{J_{<\delta}} + \mathrm{HS}_{J_{\geq\delta}} - \mathrm{HS}_{\langle 0\rangle}$ it is enough to prove the assertion for these three Hilbert series. We know that $\mathrm{HS}_{\langle 0\rangle}(t) = 1/(1-t)^{n+1}$. To compute $\mathrm{HS}_{J_{<\delta}}$ and $\mathrm{HS}_{J_{\geq\delta}}$ (Lemma 5.2) it suffices to list the monomials of degree $\leq \delta$ which are not in J. Since the number of monomials of degree $\leq \delta \leq (n+1)D - n$ is at most $(eD)^{n+1}$ and the cost of testing whether a monomial belongs to J or not is $k(n+1)$ operations, these Hilbert series is computed by $k(n+1)(eD)^{n+1}$ operations which is polynomial in D^n. □

Example 5.4. Computing the Hilbert series of a Borel type ideal. *Let R be the ring $K[x_0, x_1, x_2, x_3, x_4]$. Consider the following monomial ideal from* [4]*, Example 3.13*

$$J = \langle x_0^4, x_0^3x_1, x_0^2x_1^2, x_1^4, x_0^3x_2, x_0^2x_2^2, x_1^3x_2^5, x_0^3x_3, x_0^3x_4^2\rangle.$$

This is an ideal of dimension 3*, and it is Borel type by algorithm testing Borel type ideal. By formula* (3) *we have $\mathrm{HS}_{J_{<\delta}}(t) = P(t)/(1-t)^5$ where $\delta = 12$ and $P(t)$ is*

$$1+3t^8-6t^7+2t^6+7t^5-7t^4+t^{10}-t^{11}+1271t^{12}-4613t^{13}+6327t^{14}-3883t^{15}+899t^{16}.$$

By a simple computation (using the software [9]*) we have $a_0 = 1, a_1 = 12, a_2 = 72, a_3 = 187$ and $a_4 = 899$ (see the notation of Lemma 5.2). Thus, $\mathrm{HS}_{J_{\geq\delta}}(t)$ is equal to*

$$\frac{1 - t^{12}(13 - 12t + 72(1-t)^2 + 287(1-t)^3 + 899(1-t)^4)}{(1-t)^5}$$

which follows that

$$\begin{aligned}\mathrm{HS}_J(t) &= \mathrm{HS}_{J_{<\delta}} + \mathrm{HS}_{J_{\geq\delta}} - \frac{1}{(1-t)^5}\\ &= \frac{1 + 2t + 3t^2 + 4t^3 - 2t^4 - t^5 + 2t^6 - t^7 - t^8 - t^9}{(1-t)^3}.\end{aligned}$$

6. Degree of Borel type ideals

In this section, we give a formula for the degree of a Borel type ideal. We recall first the definition of the degree of a monomial ideal $J \subset R$.

Proposition 6.1. *We have $\mathrm{HS}_J(t) = N(t)/(1-t)^d$ where $N(t)$ is a polynomial which is not multiple of $1-t$, and $d = \dim(J)$.*

For the proof of this proposition see [8], Theorem 7 of Chapter 11. Using this proposition we define the degree of a monomial ideal.

Definition 6.2. *The degree of a monomial ideal $J \subset R$, noted by $\deg(J)$, is $N(1)$ where N is the numerator of HS_J.*

Let J be a Borel type ideal and $\delta = D_0 + \cdots + D_n - n$ where $x_0^{D_0} \cdots x_n^{D_n}$ is the least common multiple of the minimal generators of J. Using the formula $\mathrm{HS}_J = \mathrm{HS}_{J_{<\delta}} + \mathrm{HS}_{J_{\geq\delta}} - \mathrm{HS}_{\langle 0\rangle}$, and the fact that $\mathrm{HS}_{J_{<\delta}} - \mathrm{HS}_{\langle 0\rangle}$ is a polynomial in t, we can conclude that (Lemma 5.2) $N(t) = (1 - t^\delta \sum_{i=0}^{n} a_i(1-t)^i)/(1-t)^{n+1-d}$ where a_i is the number of monomial $m \in J_\delta$ such that i is the maximal integer ℓ with $x_\ell \mid m$. Let $s = 1-t$. Hence, $N(1-s) = (1-(1-s)^\delta \sum_{i=0}^{n} a_i s^i)/s^{n+1-d}$. Thus the degree of J which is $N(1)$ is the coefficient of s^{n+1-d} in $1-(1-s)^\delta \sum_{i=0}^{n} a_i s^i$ which is equal to $-\sum_{i=0}^{n-d+1}(-1)^i a_{n+1-d-i}\binom{\delta}{i}$. We summarize this result in the following theorem.

Theorem 6.3. *Let $J \subset R$ be a Borel type ideal. The degree of J is*

$$-\sum_{i=0}^{n-d+1}(-1)^i a_{n+1-d-i}\binom{\delta}{i}$$

where $d = \dim(J)$, $\delta = D_0+\cdots+D_n-n$ with $x_0^{D_0}\cdots x_n^{D_n}$ the least common multiple of the minimal generators of J and a_i is the number of monomial $m \in J_\delta$ such that i is the maximal integer ℓ with $x_\ell \mid m$.

Remark 6.4. *Since the ideal $J+\langle x_n,\ldots,x_{n-d+1}\rangle$ is zero-dimensional, then $x_i^{D_i} \in J$ for $i = 0,\ldots,n-d$. Thus, $a_0 = 1$ and $a_i = \binom{i+\delta}{i} - a_{i-1} - \cdots - a_0$ for any $i < n-d+1$.*

Corollary 6.5. *Let $J = \langle m_1,\ldots,m_k\rangle$ be a Borel type ideal. The degree of J may be computed by $k(n+1-d)\binom{n-d+\delta}{n-d}$ operations.*

From this corollary, we conclude that the complexity of computing the degree of a Borel type ideal by this theorem is sharper than computing it using Hilbert series of the ideal.

Example 6.6. Computing the degree of a Borel type ideal. *Let us consider the ideal of Example 5.4. Its degree is equal to*

$$-(72\binom{12}{0} - 12\binom{12}{1} + \binom{12}{2}) = 6.$$

7. Conclusion

In this paper, we have presented a new algorithm which computes the Hilbert series of a Borel type ideal within a complexity polynomial in D^n where $n+1$ is the number of unknowns and D is the highest degree of a minimal generator of input polynomials. We have shown also that the satiety, Castelnuovo-Mumford regularity and degree of such an ideal may be computed within the above complexity.

References

[1] D. Bayer and M. Stillman. A criterion for detecting m-regularity. *Invent. Math.*, 87(1):1–11, 1987.

[2] D. Bayer and M. Stillman. Computation of Hilbert functions. *J. Symbolic Comput.*, 14(1):31–50, 1992.

[3] I. Bermejo and P. Gimenez. Computing the Castelnuovo-Mumford regularity of some subschemes of $\mathbb{P}^n_K$ using quotients of monomial ideals. *J. Pure Appl. Algebra*, 164(1-2):23–33, 2001. Effective methods in algebraic geometry (Bath, 2000).

[4] I. Bermejo and P. Gimenez. Saturation and Catelnuovo-Mumford regularity. *J. Algebra*, 303:592–617, 2006.

[5] A. M. Bigatti, P. Conti, L. Robbiano, and C. Traverso. A "divide and conquer" algorithm for Hilbert-Poincaré series, multiplicity and dimension of monomial ideals. In *Applied algebra, algebraic algorithms and error-correcting codes (San Juan, PR, 1993)*, volume 673 of *Lecture Notes in Comput. Sci.*, pages 76–88. Springer, Berlin, 1993.

[6] D. Cox, J. Little, and D. O'Shea. *Using algebraic geometry*, volume 185 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1998.

[7] S. Eliahou and M. Kervaire. Minimal resolutions of some monomial ideals. *J. Algebra*, 129(1):1–25, 1990.

[8] R. Fröberg. *An introduction to Gröbner bases*. Pure and Applied Mathematics (New York). John Wiley & Sons Ltd., Chichester, 1997.

[9] G.-M. Greuel, G. Pfister, and H. Schönemann. SINGULAR 3.0. A Computer Algebra System for Polynomial Computations, Centre for Computer Algebra, University of Kaiserslautern, 2005. `http://www.singular.uni-kl.de`.

[10] A. Hashemi. Strong Nœther Position and Stabilized Regularities. submitted to Applicable Algebra in Engineering, Communication and Computing, 2006.

[11] A. Hashemi. Polynomial-Time Algorithm for Hilbert Series of Borel Type Ideals. In Marc Moreno Maza, Stephan M. Watt, editors, Proceedings of Symbolic-Numeric Computation (SNC 2007), pages 97-103. ACM press 2007.

[12] J. Herzog, D. Popescu, and M. Vladoiu. On the Ext-modules of ideals of Borel type. In *Commutative algebra (Grenoble/Lyon, 2001)*, volume 331 of *Contemp. Math.*, pages 171–186. Amer. Math. Soc., Providence, RI, 2003.

[13] A. Imran and A. Sarfraz. Regularity of ideals of Borel type is linearly bounded. *Preprint (see math.AC/0610537)*, 2007.

[14] E. W. Mayr and A. R. Meyer. The complexity of the word problems for commutative semigroups and polynomial ideals. *Adv. in Math.*, 46(3):305–329, 1982.

DEPARTMENT OF MATHEMATICAL SCIENCES, ISFAHAN UNIVERSITY OF TECHNOLOGY, ISFAHAN 841568G111, IRAN.

INRIA-SALSA PROJECT/ LIP6-SPIRAL TEAM, 104 AVENUE DU PRESIDENT KENNEDY, 75016 PARIS, FRANCE.

E-mail address: `Amir.Hashemi@lip6.fr`

ALBANIAN JOURNAL
OF MATHEMATICS
Volume 1, Number 3, Pages 157–172
ISSN 1930-1235: (electronic version)

SAMPLE EXTREMES OF L_p-NORM ASYMPTOTICALLY SPHERICAL DISTRIBUTIONS

ENKELEJD HASHORVA

Abstract. In this paper we deal with the asymptotic behaviour of sample maxima of L_p-norm asymptotically spherical random vectors. If the distribution function of the associated random radius of such random vectors is in the Gumbel of the Weibull max-domain of attraction we show that the normalised sample maxima has asymptotic independent components converging in distribution to a random vector with unit Gumbel or Weibull components. When the associated random radius has distribution function in the Fréchet max-domain we show that the sample maxima has asymptotic dependent components.

1. Introduction

Let $\boldsymbol{X} = (X_1, \ldots, X_d)^\top, d \geq 2$, be a random vector in $\mathbb{R}^d, d \geq 2$, defined by

$$\boldsymbol{X} = R\boldsymbol{U}_d, \tag{1}$$

where R is an almost surely positive random variable independent of the random vector $\boldsymbol{U}_d = (U_1, \ldots, U_d)^\top$ ($^\top$ stands here for the transpose sign).
Suppose that for some $p \in (0, \infty)$ we have almost surely

$$\sum_{i=1}^{d} |U_i| = 1$$

and the random vector $(U_1, \ldots, U_{d-1})^\top$ possesses probability density function

$$p(u_1, \ldots, u_{d-1}) = \frac{p^{d-1}\Gamma(d/p)}{2^{d-1}\Gamma(1/p)} \Big(1 - \sum_{i=1}^{d-1} |u_i|^p\Big)^{(1-p)/p}, \quad i \leq d-1,$$

defined for any $u_i \in [-1, 1], i \leq d$ such that $\sum_{i=1}^{d-1} |u_i|^p < 1$, where $\Gamma(\cdot)$ denotes the Gamma function.
Following Gupta and Song (1997) we shall call $\boldsymbol{X}$ with stochastic representation (1) a L_p-norm spherical random vector. In the case $p = 2$ the random vector $\boldsymbol{X}$ reduces to a spherical symmetrical (L_2-norm) random vector with the distribution function invariant with respect to orthogonal transformations in $\mathbb{R}^d$.

Received by the editors January 15, 2007 and in revised form, August 20, 2007.
2000 *Mathematics Subject Classification.* Primary 60F05; Secondary 60G70.
Key words and phrases. L_p-norm spherical random vectors, L_p-norm asymptotically spherical random vectors, asymptotic dependence, max-domain of attraction, weak convergence, multivariate regular variation.

Gupta and Song (1997), Szabłowski (1998) derive the basic distributional properties of L_p-norm spherical random vectors. The asymptotic properties of this class of random vectors have not been investigated so far, in particular in the literature no result is available for the asymptotic behaviour of sample maxima, when considering samples with underling L_p-norm spherical distributions.

As it is the case for the distributional properties, for $p = 2$ several asymptotic properties of spherical random vectors are available with some early works going back to Carnal (1970), Gale (1980), Berman (1982) among several others.
In this paper we show that the basic asymptotic properties of L_2-norm spherical random vectors extend (with minor adjustments) naturally to the general L_p-norm. With motivation from Hashorva (2005) we introduce the class of L_p-norm asymptotically spherical random vectors. We shall show that this new class of random vectors is a natural generalisation of the L_p-norm spherical random vectors with respect to the asymptotic dependence and asymptotic behaviour of sample maxima; thus generalising the results of the aforementioned paper for the L_2-norm setup.

In the next section we shall introduce some notation and provide details on L_p-norm spherical random vectors and investigate their asymptotic dependence.

We then introduce L_p-norm asymptotically spherical random vectors and discuss the main asymptotic properties of this novel class. The proofs of all the results as well as some related results are relegated to Section 5.

2. Preliminaries

We start with presenting the notation and the basic distributional properties of L_p-norm spherical random vectors. For any vector $\boldsymbol{x} = (x_1, \ldots, x_d)^\top \in \mathbb{R}^d, d \geq 2$ set $\boldsymbol{x}_I := (x_i, i \in I)^\top$ with I being a non-empty subset of $\{1, \ldots, d\}$. We shall write $\boldsymbol{x}_I^\top$ instead of $(\boldsymbol{x}_I)^\top$. Let $\boldsymbol{y} = (y_1, \ldots, y_d)^\top$ be another vector in $\mathbb{R}^d$. We define

$$\begin{gathered}
\boldsymbol{x} + \boldsymbol{y} := (x_1 + y_1, \ldots, x_d + y_d),\\
\boldsymbol{x} > \boldsymbol{y}, \text{ if } x_i > y_i, \quad \forall\, i = 1, \ldots, d,\\
\boldsymbol{x} \geq \boldsymbol{y}, \text{ if } x_i \geq y_i, \quad \forall\, i = 1, \ldots, d,\\
\boldsymbol{x} \neq \boldsymbol{y}, \text{ if for some } i \leq d \; x_i \neq y_i,\\
\boldsymbol{a}\boldsymbol{x} := (a_1 x_1, \ldots, a_d x_d)^\top, \quad c\boldsymbol{x} := (cx_1, \ldots, cx_d)^\top, \quad \boldsymbol{a} \in \mathbb{R}^d, c \in \mathbb{R},\\
\mathbf{0} := (0, \ldots, 0)^\top \in \mathbb{R}^d, \quad \mathbf{1} := (1, \ldots, 1)^\top \in \mathbb{R}^d,\\
\|\boldsymbol{x}_I\|_p := \Big(\sum_{i \in I} |x_i|^p\Big)^{1/p}, \quad \mathbb{S}_p^{k-1} := \{\boldsymbol{x} \in \mathbb{R}^k : \|\boldsymbol{x}\|_p = 1\}, \quad k \geq 1, p > 0.
\end{gathered}$$

Note that $\|\cdot\|_p$ is only for $p \in [1, \infty)$ a norm. We still refer to L_p-norm spherical random vectors even when $p \in (0, 1)$.
We shall write $Beta(a, b)$ for the distribution function of a Beta random variable with positive parameters a and b. If Z is a random variable with distribution function G we shall use alternatively the notation $Z \sim G$, and denote by G^{-1} the generalised inverse of G.

Let p be a given positive constant, and let $\boldsymbol{U}_k, k \geq 2$, denote a L_p-norm uniformly distributed random vector on $\mathbb{S}_p^{k-1}$.
Consider $\boldsymbol{X}$ as in (1) with associated random radius $R > 0$ (almost surely) with the distribution function F independent of the random vector $\boldsymbol{U}_d = (U_1, \ldots, U_d)^\top$. For $p = 2$ Cambanis et al. (1981) show that for any I, J two non-empty disjoint index

sets such that $I \cup J = \{1, \ldots, d\}$ the random vector $\boldsymbol{X}$ possesses the stochastic representation

$$\boldsymbol{X}_I \stackrel{d}{=} RW_{m,d}\boldsymbol{U}_m, \quad \text{and } \boldsymbol{X}_J \stackrel{d}{=} R(1-W_{m,d}^2)^{1/2}\boldsymbol{U}_{d-m},$$

where $\boldsymbol{U}_m, \boldsymbol{U}_{d-m}, m := |I|$ are two uniformly distributed random vectors on $\mathbb{S}_2^{m-1}$ and $\mathbb{S}_2^{d-m-1}$, respectively, and

$$W_{m,d}^2 \sim Beta(m/2,(d-m)/2), \quad W_{m,d} > 0,$$

where $\stackrel{d}{=}$ stands for equality of distribution functions of random vectors.
In Gupta and Song (1997) the above stochastic representation is generalised to the L_p-norm spherical random vectors. Referring to Theorem 3.1 therein we have for any $p > 0$ and $\boldsymbol{X}$ defined be (1)

$$\boldsymbol{X}_I \stackrel{d}{=} RW_{m,d,p}\boldsymbol{U}_m, \quad \text{and } \boldsymbol{X}_J \stackrel{d}{=} R(1-W_{m,d,p}^p)^{1/p}\boldsymbol{U}_{d-m}, \tag{2}$$

with $\boldsymbol{U}_m, \boldsymbol{U}_{d-m}$ two independent L_p-norm uniformly distributed random vectors. Further, $R, W_{m,d,p}, \boldsymbol{U}_m, \boldsymbol{U}_{d-m}$ are mutually independent, and

$$W_{m,d,p}^p \sim Beta(m/p,(d-m)/p), \quad W_{m,d,p} > 0.$$

As shown initially by Berman (1992) the stochastic representation (2) is basic for investigating the asymptotic behaviour of L_2-norm spherical random vectors.
Utilising Berman's results and the ideas Hashorva (2005) discusses the asymptotic dependence and the asymptotic behaviour of sample extremes of asymptotically spherical and elliptical random vectors.
Next we shall consider the asymptotic dependence of L_p-norm spherical random vectors.
In view of the stochastic representation (2) we simply need to investigate the asymptotic dependence of a bivariate L_p-norm spherical random vector.
Let therefore $p > 0$ be fixed and let $\boldsymbol{X} = (X_1, X_2)^\top$ be a L_p-norm bivariate spherical random vector. By (2) both X_1, X_2 have the same distribution and are symmetric about 0. Denote by F the distribution function of the associated random radius R. The simple (well-known) measure of asymptotic dependence between X_1, X_2 is the limit (if it exists)

$$\kappa(X_1, X_2) := \lim_{t \uparrow \omega} \frac{\boldsymbol{P}\{X_1 > t, X_2 > t\}}{\boldsymbol{P}\{X_1 > t\}} \geq 0,$$

with $\omega := \sup\{x : F(x) < 1\}$ the upper endpoint of F. If $\kappa(X_1, X_2) = 0$ then the joint tail probability diminishes faster than the marginal tail probability. For this case we say that X_1 and X_2 are asymptotically independent.

If ω is finite, then $\kappa(X_1, X_2) = 0$ since both X_1, X_2 cannot approach ω simultaneously. We discuss in the following therefore only the case $\omega = \infty$.
Let $a \in (0, \infty)$ and set $c_1^p := \inf\{|x_1|^p + |x_2|^p : x_1 \geq 1, x_2 \geq a\}$. Clearly, c_1 exists and $c_1 > 1$. For any $t > 0, c_0 \in (0,1)$ we have

$$\begin{aligned} &\frac{\boldsymbol{P}\{X_1 > t, X_2 > at\}}{\boldsymbol{P}\{X_1 > t\}} \\ &\leq \frac{\boldsymbol{P}\{|X_1|^p + |X_2|^p \geq c_1^p t^p\}}{\boldsymbol{P}\{X_1 > t\}} = \frac{2\boldsymbol{P}\{R \geq c_1 t\}}{\boldsymbol{P}\{|X_1| > t\}} \\ &\leq \frac{2\boldsymbol{P}\{R \geq c_1 t\}}{\boldsymbol{P}\{RW_{1,2,p} > t, W_{1,2,p} < c_0\}} \leq \frac{2\boldsymbol{P}\{R \geq c_1 t\}}{\boldsymbol{P}\{R > t\}\boldsymbol{P}\{W_{1,2,p} < c_0\}}. \end{aligned} \tag{3}$$

Choosing a $c_0 \in (c_1^{-1}, 1)$ we obtain thus by the above upper bound

$$\kappa(X_1, X_2/a) = \lim_{t\to\infty} \frac{\boldsymbol{P}\{X_1 > t, X_2 > at\}}{\boldsymbol{P}\{X_1 > t\}} \quad = \quad 0, \tag{4}$$

provided that

$$\lim_{t\to\infty} \frac{1 - F(Kt)}{1 - F(t)} = 0 \tag{5}$$

holds for any $K > 1$. (5) means that $1 - F$ is a rapidly varying function. See de Haan (1970) or Resnick (1987) for the main properties of rapidly varying functions. In particular (5) holds if F is in the max-domain of attraction of the unit Gumbel distribution function $\Lambda(x) = \exp(-\exp(-x)), x \in \mathbb{R}$. A necessary and sufficient condition for F to be in the max-domain of attraction of Λ is the existence of a positive scaling function w such that

$$\lim_{t\uparrow\omega} \frac{1 - F(t + x/w(t))}{1 - F(t)} = \exp(-x), \quad \forall x \geq 0. \tag{6}$$

See Leadbetter et al. (1983), Galambos (1987), Resnick (1987), Reiss (1989), Berman (1992), or Falk et al. (2004) for further details.

It follows from the univariate extreme value theory that the two other possible max-domain of attractions for F are the Weibull and the Fréchet ones. For the first case we have for some $\alpha > 0$

$$\lim_{t\to\infty} F^t(\omega + a(t)x) = \exp(-|x|^\alpha) =: \Psi_\alpha(x), \quad \forall x \in (-\infty, 0), \tag{7}$$

with $a(t) := \omega - F^{-1}(1 - 1/t), t > 1$, whereas for the second case

$$\lim_{t\to\infty} F^t(a(t)x) = \exp(-x^{-\alpha}) =: \Phi_\alpha(x), \quad \forall x \in (0, \infty) \tag{8}$$

holds with $a(t) := F^{-1}(1 - 1/t), t > 1$.

If F is in the Weibull max-domain of attraction, then necessarily $\omega < \infty$, consequently (4) holds for any p positive.

Thus in both Gumbel and Weibull cases asymptotic independence of the components is observed for L_p-norm spherical random vectors. Berman (1992) shows that for F in the Fréchet max-domain of attraction $\kappa(X_1, X_2/a)$ is positive for any $a > 0$. It is well-known (see e.g. de Haan (1970) or Kotz and Nadarajah (2005)) that (8) is equivalent with the fact that R is regularly varying with index $\alpha > 0$, i.e.

$$\lim_{t\to\infty} \frac{\boldsymbol{P}\{R > Kt\}}{\boldsymbol{P}\{R > t\}} = K^{-\alpha}, \quad \forall K > 0.$$

Berman (1992) shows further for the case $p = 2$ (see Theorem 12.3.2 therein and Theorem 5.1 below) that also $|X_1|$ is regularly varying with positive index $\alpha > 0$. In Hashorva (2006) (see also Hashorva (2007b)) the converse is proved, i.e. if $|X_1|$ is regularly varying then the associated random radius R is also regularly varying with the same index as $|X_1|$.

We show in the next section that a similar result holds for L_p-norm spherical random vectors with $p > 0$ a given constant. In particular we have

$$\kappa(X_1, X_2/a) > 0, \quad \forall a \in (0, \infty)$$

if X_1 or R is regularly varying with positive index α.

3. Asymptotics of Sample Maxima

Let $\boldsymbol{X}$ be a L_p-norm spherical random vector in $\mathbb{R}^d, d \geq 2$ as in (1) and let further $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n, n \geq 1$ be independent random vectors in $\mathbb{R}^d$ with the same distribution function G as $\boldsymbol{X}$. Denote by F the distribution function of the associated random radius R of $\boldsymbol{X}$, and define the component-wise sample maxima by

$$\boldsymbol{M}_n := (\max_{1\leq j\leq n} X_{j1}, \ldots, \max_{1\leq j\leq n} X_{jd})^\top, \quad n \geq 1.$$

Assuming that F is in the max-domain of attraction of a univariate extreme value distribution function H (abbreviated as $F \in MDA(H)$) Hashorva (2005) derives (when $p = 2$) the convergence in distribution

$$\frac{\boldsymbol{M}_n - b(n)\mathbf{1}}{a(n)} \quad \xrightarrow{d} \quad \boldsymbol{Z}, \quad n \to \infty, \tag{9}$$

where the random vector $\boldsymbol{Z} = (Z_1, \ldots, Z_d)^\top$ has distribution function Q which is a product distribution if $H = \Lambda$ or $H = \Psi_\alpha, \alpha > 0$.
If $H = \Phi_\alpha, \alpha > 0$, then $\boldsymbol{Z}$ has dependent components with distribution function Φ_α. Both constants $a(n) > 0, b(n), n \in \mathbb{N}$ are defined in terms of the distribution function of X_1. The convergence in the distribution in (9) is equivalent with

$$\lim_{n\to\infty} G^n(a(n)\boldsymbol{x} + b(n)\mathbf{1}) = Q(\boldsymbol{x}), \quad \forall \boldsymbol{x} \in \mathbb{R}^d. \tag{10}$$

As in the univariate setup abbreviate (10) by $G \in MDA(Q)$. Note in passing that (10) implies $G_i \in MDA(Q_i), 1 \leq i \leq d$ with G_i, Q_i the marginal distributions of G and Q, respectively.

In the next two theorems we show that the asymptotic behaviour of the sample maxima is the same (with respect to the limiting distribution Q) for any $p > 0$. We discuss first the case F is in the Gumbel or the Weibull max-domain of attractions.

Theorem 1. *Let $\boldsymbol{X}$ be a L_p-norm spherical random vector in $\mathbb{R}^d, d \geq 2$ with distribution function G defined in (1) with $R > 0$ almost surely being independent of $\boldsymbol{U}_d$. Let F be the distribution function of R with the upper endpoint $\omega \in (0, \infty]$.*
i) Assume that $F \in MDA(\Lambda)$ with positive scaling function w. Then (10) holds where $\boldsymbol{Z}$ has independent components with unit Gumbel distribution and $b(n) := G_1^{-1}(1 - 1/n), a(n) := 1/w(b(n)), n > 1$.
ii) Suppose that $\omega = 1$ and further $F \in MDA(\Psi_\alpha), \alpha > 0$ holds. Then (10) holds with $a(n) := 1 - G_1^{-1}(1 - 1/n), n > 1$ and $Z_i, 1 \leq i \leq d$ independent random variables such that $Z_i \sim \Psi_{\alpha+(d-1)/p}$.

Several examples may illustrate the applicability of Theorem 1.

Example 1. [L_p-norm Kotz Type I] Let $\boldsymbol{X} = (X_1, \ldots, X_d)^\top$ be a random vector in $\mathbb{R}^d, d \geq 2$, with density function given by

$$q(\boldsymbol{x}) := \frac{p^d \Gamma(d/p) r^{(d/p+N)/s} s}{2^d \Gamma^d(1/p) \Gamma((d/p + N)/s)} \|\boldsymbol{x}\|_p^{pN} \exp(-r\|\boldsymbol{x}\|_p^{ps}), \quad \boldsymbol{x} \in \mathbb{R}^d \tag{11}$$

and constants $p, r, s > 0, N \in \mathbb{R} : d + pN > 0$. We refer to $\boldsymbol{X}$ as L_p-norm Kotz Type I random vector. It has stochastic representation (1) where the associated random radius R possesses the density function

$$f(t) \quad = \quad \frac{psr^{(d/p+N)/s}}{\Gamma((d/p + N)/s)} t^{d+pN-1} \exp(-rt^{ps}), \quad t \in (0, \infty).$$

If $s=1, N=0, r>0$, then $\boldsymbol{X}$ possesses a p-generalised Gaussian distribution (see Gordon and Kotz (1973)). It now follows easily that the distribution function F of the density f is in the Gumbel max-domain of attraction with the scaling function

$$w(u) \quad = \quad (1+o(1))rpsu^{ps-1}, \quad u \to \infty.$$

Next, if $\boldsymbol{X}_n, n \geq 1$ are independent with density function q given in (11), then (9) follows implying further that the sample maxima has asymptotic independent components.

Example 2. [L_p-norm Kotz Type III] We call a random vector $\boldsymbol{X}$ in $\mathbb{R}^d, d \geq 2$ a L_p-norm Kotz Type III spherical random vector if it has stochastic representation (1) where the associated random radius $R>0$ has asymptotic tail behaviour ($u \to \infty$)

$$\boldsymbol{P}\{R>u\} \quad = \quad (1+o(1))Ku^N \exp(-ru^\delta), \quad K>0, \delta>0, N \in \mathbb{R}. \tag{12}$$

It now easily follows that the distribution function F of R is in the Gumbel max-domain of attraction with the positive scaling function

$$w(u) \quad = \quad (1+o(1))r\delta u^{\delta-1}, \quad u \to \infty.$$

The subvectors of $\boldsymbol{X}$ are all L_p-norm Kotz Type III spherical random vectors. This property is not shared by L_p-norm Kotz Type I spherical random vectors.
In view of Theorem 1 the random vector $\boldsymbol{X}$ has asymptotic independent components, and the maxima of a sample of L_p-norm Kotz Type III spherical random vectors has asymptotic independent components with distribution function attracted by a product distribution with Gumbel marginals.

Example 3. [L_p-norm Pearson Type II] The random vector $\boldsymbol{X}$ in $\mathbb{R}^d, d \geq 2$ has density function (see Example 2.3 of Gupta and Song (1997))

$$q(\boldsymbol{x}) \quad := \quad \frac{p^d \Gamma(d/p+\alpha)}{2^d \Gamma^d(1/p)\Gamma(\alpha)} \Big(1-\|\boldsymbol{x}\|_p^p\Big)^{\alpha-1}, \quad \boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x}\|_p < 1, \alpha>0.$$

$\boldsymbol{X}$ is L_p-norm spherically distributed with random radius R with density function

$$f(t) \quad = \quad \frac{p\Gamma(d/p+\alpha)}{\Gamma(d/p)\Gamma(\alpha)} t^{d-1}(1-t^p)^{\alpha-1}, \quad t \in (0,1).$$

It follows that the associated random radius R has the distribution function in the max-domain of attraction of the Weibull distribution Ψ_α. Hence by the above theorem the distribution function of $\boldsymbol{X}$ is in the max-domain of attraction of a product distribution with marginal distributions $\Psi_{\alpha+(d-1)/p}$.

Next, we deal with the case where F is in Fréchet max-domain of attraction. In Hashorva (2005) (see also Hashorva (2007b)) it is shown (considering only the case $p=2$) that R has distribution function $F \in MDA(\Phi_\alpha), \alpha>0$ iff X_1 has distribution function in the max-domain of attraction of Φ_α.
Further, it is proved therein that $F \in MDA(\Phi_\alpha)$ implies that $\boldsymbol{X}$ is a regularly varying random vector with index α. Regular variation of random vectors is investigated in details in many recent contributions, see e.g. Basrak (2002), Mikosch (2005). We use the following definitions of regular variation of random vectors.

Definition 1. The random vector $\boldsymbol{X}=(X_1,\ldots,X_d)^\top, d \geq 1$ is regularly varying with index $\alpha>0$ if there exists a positive sequence $a_n \to \infty$ as $n \to \infty$ and a

positive measure μ homogeneous of order α such that the vague convergence in $[-\infty,\infty]^d \setminus \{\mathbf{0}\}$

$$n\boldsymbol{P}\{\boldsymbol{X}/a_n \in \cdot\} \quad \stackrel{v}{\to} \quad \mu(\cdot), \quad n \to \infty \tag{13}$$

holds.

For $\boldsymbol{X}$ a spherical random vector (with respect to L_2-norm) in $\mathbb{R}^d, d \geq 2$ Hashorva (2006) shows that if X_1 is regularly varying with index α then $\boldsymbol{X}$ is regularly varying with the same index α.

We generalise the aforementioned results for the case $p > 0$ (and complete the proof of our previous result). The asymptotic behaviour of the sample maxima is derived in Corollary 3.

Theorem 2. *Let $\boldsymbol{X}$ be as in Theorem 1, and let $R_{i,p} := (\sum_{j=1}^{i}|X_j|^p)^{1/p}, 1 \leq i \leq d$ be the i-th associated random radius of $\boldsymbol{X}$. Then the following statements are equivalent:*
i) $R_{d,p}$ is regularly varying with index $\alpha > 0$.
ii) X_1 is regularly varying with index $\alpha > 0$.
iii) For any $i = 1,\ldots,d$ the random radius $R_{i,p}$ is regularly varying with positive index α. Furthermore

$$\boldsymbol{P}\{R_{i,p} > u\} \quad = \quad (1+o(1))C_{i,d,\alpha,p}\boldsymbol{P}\{R_{d,p} > u\}, \quad u \to \infty \tag{14}$$

holds for any $i < d$ where

$$C_{i,d,\alpha,p} \quad := \quad 2\frac{\Gamma(d/p)\Gamma((i+\alpha)/p)}{\Gamma(i/p)\Gamma((d+\alpha)/p)} \in (0,\infty). \tag{15}$$

iv) For any $I \subset \{1,\ldots,d\}$ with $1 \leq |I| = k$ and any Borel set B away from the origin of $\mathbb{R}^k$

$$\lim_{u\to\infty}\frac{\boldsymbol{P}\{(\boldsymbol{X}+\boldsymbol{\mu})_I/u \in B\}}{\boldsymbol{P}\{X_1 > u\}} \quad = \quad C_{1,k,\alpha,p}\int_0^\infty \boldsymbol{P}\{r\boldsymbol{U}_k \in B\}\, d(r^{-\alpha}) \tag{16}$$

holds for any $\boldsymbol{\mu} \in \mathbb{R}^d$.
v) $\boldsymbol{X}$ is regularly varying with index $\alpha > 0$.

We thus immediately obtain the corollary:

Corollary 3. *Let $\boldsymbol{X}, \boldsymbol{U}_d, d \geq 2$ be as in Theorem 1 with distribution function G. Assume that R or X_1 is regularly varying with positive index α. Then $G \in MDA(Q_{d,\alpha,p})$ with $Q_{d,\alpha,p}$ a max-stable distribution function on $(0,\infty)^d$ defined for any $\boldsymbol{x} > \mathbf{0}$ by*

$$Q_{d,\alpha,p}(\boldsymbol{x}) \quad := \quad \exp\Big(-C_{1,d,\alpha,p}\int_0^\infty \boldsymbol{P}\{r\boldsymbol{U}_d \in \mathbb{R}^d \setminus \times_{i=1}^d(-\infty,x_i]\}\, d(r^{-\alpha})\Big). \tag{17}$$

The marginal distributions of $Q_{d,\alpha,p}$ are identical to Φ_α.
Conversely, if $G \in MDA(Q_{d,\alpha,p})$ where $Q_{d,\alpha,p}$ has marginal distributions identical to $\Phi_\alpha, \alpha > 0$, then both X_1 and R_1 are regularly varying with index α.

(17) implies that if $\boldsymbol{Y}$ is a random vector with distribution function $Q_{d,\alpha,p}$ for some $d \geq 2$ and α, p positive constants, then the subvector $\boldsymbol{X}_I$ where I has $k \geq 1$ elements and $I \subset \{1,\ldots,d\}$ has distribution function $Q_{k,\alpha,p}$ which is max-stable. Furthermore, $Q_{k,\alpha,p}$ is not a product distribution for any $k \geq 2$.

We present next an illustrating example:

Example 4. [L_p-norm Pearson Type VII] Define $\boldsymbol{X}$ a L_p-norm spherically distributed random vector in $\mathbb{R}^d, d \geq 2$ as in Example 2.4 of Gupta and Song (1997) with density function given for any $\boldsymbol{x} \in \mathbb{R}^d$ by

$$q(\boldsymbol{x}) \quad := \quad \frac{p^d \Gamma(N)}{2^d \Gamma^d(1/p)\Gamma(N-n/p)} s^{-d/p} \Big(1 + \|\boldsymbol{x}\|_p^p/s\Big)^{-N}, \quad s > 0, N > d/p.$$

If $N = (d+m)/2$ then $\boldsymbol{X}$ has a L_p-norm t-distribution (see Example 2.5 of Gupta and Song (1997)). The associated random radius R has density function

$$f(t) \quad = \quad \frac{p\Gamma(N)}{\Gamma(d/p)\Gamma(N-d/p)} s^{-d/p} t^{d-1} \Big(1 + t^p/s\Big)^{-N}, \quad t \in (0, \infty).$$

In view of Karamata's Theorem (see e.g. Resnick (1987)) the random variable R is regularly varying with index $\alpha := pN - d > 0$. Consequently the marginals $X_i, 1 \leq i \leq d$ are regularly varying with index α. Furthermore, $\boldsymbol{X}$ is a regularly varying random vector with positive index α. The corresponding measure can be easily calculated.

Example 5. [L_p-norm Kotz Type II] We say that a random vector $\boldsymbol{X}$ in $\mathbb{R}^d, d \geq 2$, has L_p-norm Kotz Type II distribution if its density function is given by

$$q(\boldsymbol{x}) := \frac{p^d \Gamma(d/p) r^{d/p+N} s}{2^d \Gamma^d(1/p)\Gamma((d/p+N)/s)} \|\boldsymbol{x}\|_p^{pN} \exp(-r\|\boldsymbol{x}\|_p^{ps}), \quad \boldsymbol{x} \in \mathbb{R}^d, \tag{18}$$

with constants $p > 0, r > 0, s < 0, d/p + N < 0$. Kotz (1975) introduces $\boldsymbol{X}$ with density function as above in the case $p = 2$. Basic properties of $A\boldsymbol{X}$ with $A \in \mathbb{R}^{d\times d}$ a non-singular matrix are discussed in Kotz (2004). It can easily be shown that $\boldsymbol{X}$ has stochastic representation (1) with the random radius R which has distribution function in the Fréchet max-domain of attraction. Consequently, Theorem 2 implies that the components of $\boldsymbol{X}$ are asymptotically dependent and the sample maxima of Kotz Type II random vectors converges in the distribution (after normalisation) to a random vector with dependent Fréchet marginal components.

4. L_p-norm Asymptotically Spherical Random Vectors

L_2-norm asymptotically spherical random vectors are introduced in Hashorva (2005). The crucial asymptotic property of such vectors is that the asymptotic behaviour of the sample extremes can be defined by the asymptotic behaviour of the associated random radius $R_{i,2}, i \leq d$. In this section we introduce the larger class of L_p-norm asymptotically spherical random vectors and show that the asymptotic properties of L_p-norm spherical random vectors still hold under such a general setup.

Definition 2. [L_p-norm asymptotically spherical random vector] Let $\boldsymbol{X}$ be a random vector in $\mathbb{R}^d, d \geq 2$ and let $\omega \in (0, \infty]$ be the upper endpoint of the distribution function of each component $X_i, 1 \leq i \leq d$. For any non-empty subset $I \subset \{1, \ldots, d\}$ set $R_{I,p} := (\sum_{i\in I} |X_i|^p)^{1/p}, p > 0$ and $R := (\sum_{i=1}^d |X_i|^p)^{1/p}$. Assume that $R > 0$ almost surely. If additionally

$$\lim_{t\uparrow\omega} \frac{\boldsymbol{P}\{X_i > t\}}{\boldsymbol{P}\{|X_i| > t\}} \quad = \quad c_i \in (0, 1] \tag{19}$$

and further for any non-empty index set $I \subset \{1,\dots,d\}$

$$\lim_{t\uparrow\omega} \frac{\boldsymbol{P}\{RW_{I,p} > t\}}{\boldsymbol{P}\{R_{I,p} > t\}} \quad = \quad d_I \in (0,\infty), \tag{20}$$

where $W_{I,p}^p \sim Beta(\delta_I, \lambda_I), W_{I,p} > 0$ being further independent of R, with δ_I, λ_I positive, then we refer to $\boldsymbol{X}$ as a L_p-norm asymptotically spherical random vector.

In the following we consider for simplicity the case

$$d_I = 1, \quad \forall I \subset \{1,\dots,d\}.$$

Further we suppose that for two non-empty index sets I, J such that $J \subset I \subset \{1,\dots,d\}$ we have $\delta_I \geq \delta_J$. For notational simplicity we shall write δ_i, λ_i when $I = \{i\}$.

With the above restrictions we call $\boldsymbol{X}$ a L_p-norm asymptotically spherical random vector with coefficients $\boldsymbol{c}, \delta_I, \lambda_I, I \subset \{1,\dots,d\}$ where $\boldsymbol{c} = (c_1,\dots,c_k)^\top \in \mathbb{R}^k$, or shortly a L_p-norm asymptotically spherical random vector.
If $p = 2$ an instance of L_2-norm asymptotically spherical random vector is $\boldsymbol{X}$ a generalised symmetrised Dirichlet random vector introduced in Fang and Fang (1990). The main asymptotic properties derived above for the L_p-norm spherical random vectors can be extended for the more general case L_p-norm asymptotically spherical random vectors. Since both X_1, X_2 and the associated random radius R have by definition the same upper endpoint ω, then $\kappa(X_1, X_2) = 0$ follows in the case $\omega \in (0,\infty)$. We discuss next the case $\omega = \infty$ and R has a rapidly varying survival function.

Theorem 4. *Let $\boldsymbol{X}$ be a L_p-norm asymptotically spherical random vector in $\mathbb{R}^d, d \geq 2$ with coefficients $\boldsymbol{c}, \delta_I, \lambda_I, I \subset \{1,\dots,d\}$. Let F be the distribution function of $R := (|X_1|^p + \cdots + |X_d|^p)^{1/p}$ with the upper endpoint $\omega \in (0,\infty]$.*
i) If $\omega \in (0,\infty)$ then $\kappa(X_i, X_j) = 0, 1 \leq i < j \leq d$.
ii) If $\omega = \infty$ and F is rapidly varying then for any $a > 0$ we have $\kappa(X_i, X_j/a) = 0, 1 \leq i < j \leq d$.

Let $\boldsymbol{X}$ be a L_p-norm asymptotically spherical random vector with associated random radius R. If the distribution function of R is in the max-domain of attraction of a univariate extreme value distribution, then Theorem 8 below implies that the components of $\boldsymbol{X}$ have distribution function in the same max-domain of attraction. We show next that also the distribution function of $\boldsymbol{X}$ is in the max-domain of attraction of a max-stable distribution function.

Theorem 5. *Let $\boldsymbol{X}, \boldsymbol{X}_n, n \geq 1$ be independent L_p-norm asymptotically spherical random vectors in $\mathbb{R}^d, d \geq 2$ with coefficients $\boldsymbol{c}, \delta_I, \lambda_I, I \subset \{1,\dots,d\}$ and distribution function G. Denote by $\omega \in (0,\infty]$ the upper endpoint of the distribution function F of the associated random radius $R > 0$ of $\boldsymbol{X}$.*
i) Let $\boldsymbol{Z}$ be a $d-$dimensional random vector with independent unit Gumbel components. If $F \in MDA(\Lambda)$ with the scaling function w, then we have the convergence in the distribution

$$\frac{\boldsymbol{M}_n - \boldsymbol{b}(n)}{\boldsymbol{a}(n)} \quad \xrightarrow{d} \quad \boldsymbol{Z}, \quad n \to \infty, \tag{21}$$

provided that $\lambda_i = \lambda > 0, 1 \leq i \leq d$ if $\omega = \infty$, where $\boldsymbol{a}(n), \boldsymbol{b}(n)$ are defined by

$$b_i(n) := G_i^{-1}(1 - 1/n), \quad a_i(n) := 1/w(b_i(n)), \quad 1 \leq i \leq d, n > 1.$$

ii) Assume that $\omega = 1$ *and* $F \in MDA(\Psi_\alpha), \alpha > 0$. *Then we have*

$$\frac{\boldsymbol{M}_n - \mathbf{1}}{\boldsymbol{a}(n)} \quad \stackrel{d}{\to} \quad \boldsymbol{Z}, \quad n \to \infty, \tag{22}$$

where $Z_i, 1 \le i \le d$ *are independent with* $Z_i \sim \Psi_{\alpha+\lambda_i}$, *and* $\boldsymbol{a}(n)$ *has components* $a_i(n) := 1 - G_i^{-1}(1 - 1/n), n > 1, 1 \le i \le d$.

We note in passing that the restriction $\lambda_i = \lambda > 0$ for all $i \le d$ in the above theorem might be redundant. This is the case for instance if $\boldsymbol{X}$ is a L_2-norm generalised symmetrised Dirichlet distribution.
We consider next the case that the associated random radius R is regularly varying.

Theorem 6. *Let* $\boldsymbol{X}$ *be a* L_p*-norm asymptotically spherical random vector as in Theorem 5. If the associated random radius* R *is regularly varying with positive index* α, *then* $X_i, 1 \le i \le d, R_{I,p}, \forall I \subset \{1, \ldots, d\}$ *are regularly varying with index* α *and furthermore*

$$\boldsymbol{P}\{R_{I,p} > u\} \quad = \quad (1 + o(1)) \frac{\Gamma(\delta_I + \lambda_I)\Gamma(\alpha/p + \delta_I)}{\Gamma(\delta_I)\Gamma(\alpha/p + \delta_I + \lambda_I)} \boldsymbol{P}\{R > u\} \tag{23}$$

holds as $u \to \infty$. *Furthermore, if there exists a random vector* $\boldsymbol{U}$ *on* $\mathbb{S}_p^{d-1}$ *independent of* R *such that* $\boldsymbol{X} = R\boldsymbol{U}$, *then we have for any Borel set* $B \subset \mathbb{R}^d$ *away from the origin of* $\mathbb{R}^d$ *and for any vector* $\boldsymbol{\mu} \in \mathbb{R}^d$

$$\lim_{u \to \infty} \frac{\boldsymbol{P}\{(\boldsymbol{X} + \boldsymbol{\mu})/u \in B\}}{\boldsymbol{P}\{X_i > u\}} \quad = \quad C_i \int_0^\infty \boldsymbol{P}\{r\boldsymbol{U} \in B\}\, d(r^{-\alpha}), \tag{24}$$

where

$$C_i := \frac{\Gamma(\delta_i)\Gamma(\alpha/p + \delta_i + \lambda_i)}{c_i \Gamma(\delta_i + \lambda_i)\Gamma(\alpha/p + \lambda_i)} \in (0, \infty).$$

Similarly as in Corollary 3 the distribution function G of $\boldsymbol{X}$ in the above theorem is in the max-domain of attraction of max-stable distribution function with Fréchet marginal distributions which is not a product distribution.

We conclude this section with two illustrating example.

Example 6. [L_2-norm Kotz Type I generalised symmetrised Dirichlet] Let $\boldsymbol{\alpha}$ be a fixed vector in $\mathbb{R}^d, d \ge 2$ with positive components and let N, r, s be positive constants. We refer to a random vector $\boldsymbol{X}$ in $\mathbb{R}^d$ as Kotz Type I generalised symmetrised Dirichlet with parameters $\boldsymbol{\alpha} \in (0, \infty)^d, N \in \mathbb{R}, r > 0, s > 0$ if it possesses the density function

$$h(\boldsymbol{x}) := \frac{r^{(N + \sum_{i \le d} \alpha_i)/s}}{s\Gamma((N + \sum_{i \le d} \alpha_i)/s)} \frac{\Gamma(\sum_{i \le d} \alpha_i)}{\prod_{i=1}^k \Gamma(\alpha_i)} \|\boldsymbol{x}\|_p^{pN} \exp(-r\|\boldsymbol{x}\|^{2s}) \prod_{i=1}^k |x_i|^{2\alpha_i - 1}$$

defined for all $\boldsymbol{x} \in \mathbb{R}^d, d \ge 2$.
In view of the amalgamation property shown in Fang and Fang (1990) (see also Hashorva et al. (2007b)) it follows that $\boldsymbol{X}$ is a L_2-norm asymptotically spherical random vector. The associated random radius R of $\boldsymbol{X}$ is almost surely positive and moreover R^2 is Gamma distributed with parameters $\sum_{i \le d} \alpha_i$ and $1/2$. Hence the distribution function of $\boldsymbol{X}$ is in the max-domain of attraction of a product distribution with marginal distributions Λ.

Example 7. [L_p-norm Kotz Type III asymptotically spherical] Let $\boldsymbol{X}$ be a L_p-norm asymptotically spherical random vector in $\mathbb{R}^d, d \geq 2$, with coefficients $\delta_I, \lambda_I, I \subset \{1, \ldots, d\}$ and distribution function G. We say that $\boldsymbol{X}$ is a L_p-norm Kotz Type III asymptotically spherical random vector if the associated random radius R has asymptotic tail behaviour given by (12). In view of Example 2 and Theorem 5 $G \in MDA(Q)$ with Q a product distribution with unit Gumbel marginal distributions.

5. Related Results and Proofs

The next lemma is presented in the published paper Hashorva et al. (2007) referring to this paper. We give it here for reference purposes.

Lemma 7. *Let X, Y be two independent positive random variables with $Y^p \sim Gamma(a, \lambda), a, \lambda > 0, p > 0$. If X is regularly varying with positive index γ, then we have*

$$\lim_{u\to\infty} \frac{\boldsymbol{P}\{XY > u\}}{\boldsymbol{P}\{Y > u\}} = \frac{\Gamma(a+\gamma/p)}{\lambda^{\gamma/p}\Gamma(a)} \in (0, \infty). \tag{25}$$

Conversely, if the product XY is regularly varying with index $\gamma > 0$, then X is regularly varying with index γ and further (25) *holds.*

Proof of Lemma 7 By Breiman's Lemma (see Breiman (1965)) we have

$$\lim_{u\to\infty} \frac{\boldsymbol{P}\{XY > u\}}{\boldsymbol{P}\{Y > u\}} = \frac{\Gamma(a+\gamma/p)}{\lambda^{\gamma/p}\Gamma(a)} \in (0, \infty).$$

Since X is regularly varying then the first claim follows. We show next the converse. Assume that XY is regularly varying with index $\gamma > 0$. Since $XY = (X^pY^p)^{1/p}, p > 0$ and the fact that X^p is regularly varying iff X is regularly varying it suffices to show the proof for the case $p = 1$. Next, suppose for simplicity that $p = 1, \lambda = 1$. For any $t > 0$ we may write by the independence of X and Y

$$\boldsymbol{P}\{XY > t\} = t^a \int_0^\infty \exp(-tv)\, dG(v),$$

where

$$G(s) := \frac{1}{\Gamma(a)} \int_0^s \boldsymbol{P}\{X > 1/x\} x^{a-1}\, dx, \quad s > 0.$$

The assumption XY is regularly varying with index $\gamma > 0$ means

$$\int_0^\infty \exp(-tv)\, dG(v) = t^{-a-\gamma} L(1/t), \quad t \to \infty, \tag{26}$$

with $L(x)$ such that $\lim_{t\to 0} L(Kt)/L(t) = 1, \forall K > 0$.
In view of Karamata's Tauberian Theorem (Feller (1966), Resnick (1987)) (26) is equivalent with

$$G(t) = \frac{1}{\Gamma(a+\gamma+1)} t^{a+\gamma} L(t), \quad t \downarrow 0,$$

or equivalently

$$G(1/t) = \frac{1}{\Gamma(a+\gamma+1)} t^{-a-\gamma} L(1/t), \quad t \to \infty.$$

Consequently

$$\int_0^{1/t} \boldsymbol{P}\{X > 1/x\} x^{a-1}\, dx = \frac{\Gamma(a)}{\Gamma(a+\gamma+1)} t^{-a-\gamma} L(1/t), \quad t \to \infty.$$

Since $\boldsymbol{P}\{X > x\}x^{-a-1}, x > 0$ decreases monotonically in x for any $a > 0$ we obtain applying the Monotone Density Theorem (Resnick (1987))

$$\boldsymbol{P}\{X > t\}t^{-a-1} = \frac{(a+\gamma+1)\Gamma(a)}{\Gamma(a+\gamma+1)} t^{-a-\gamma-1} L(1/t), \quad t \to \infty,$$

thus the proof follows. □

Theorem 8. *Let Y be a random variable with distribution function H which has the upper endpoint $\omega \in (0, \infty]$ and $H(0) = 0$. Let a, b, τ be positive constants and let $Z_{a,b}$ be a Beta distributed random variable with parameters a, b independent of Y and set $\overline{H}(u) := 1 - H(u), u > 0$.*
i) If $H \in MDA(\Lambda)$ with positive scaling function w then we have as $u \uparrow \omega$

$$\boldsymbol{P}\{Y[1 - Z_{a,b}]^{1/\tau} > u\} = (1 + o(1))\frac{\Gamma(a+b)}{\Gamma(b)} \left(\frac{\tau}{uw(u)}\right)^a \overline{H}(u). \tag{27}$$

ii) If $H \in MDA(\Phi_\alpha), \alpha > 0$ then $\omega = \infty$ and for $u \to \infty$

$$\boldsymbol{P}\{Y[1 - Z_{a,b}]^{1/\tau} > u\} = (1 + o(1))\frac{\Gamma(a+b)\Gamma(b+\alpha/\tau)}{\Gamma(b)\Gamma(a+b+\alpha/\tau)} \overline{H}(u). \tag{28}$$

iii) If $H \in MDA(\Psi_\alpha), \alpha > 0$ and $\omega = 1$, then we have

$$\boldsymbol{P}\{Y[1 - Z_{a,b}]^{1/\tau} > u\} = (1 + o(1))\frac{\Gamma(\alpha+1)\Gamma(a+b)}{\Gamma(b)\Gamma(\alpha+a+1)} (\tau(1-u))^a \overline{H}(u) \tag{29}$$

as $u \to \infty$.

Proof. The proof can be established along the lines of the proof of Theorem 12.3.1, Theorem 12.3.2 and Theorem 12.3.3 of Berman (1992). We give below the sketch of a slightly different proof.
Let $B(y, a, b), y \in [0, 1]$ denote the distribution function of $Z_{a,b}$ and put

$$H_u(s) := H(u + s/w(u))/\overline{H}(u), \quad \overline{H}(u) := 1 - H(u), \quad u \in \mathbb{R}, s > 0.$$

Since $Y > 0$ is independent of $Z_{a,b}$ we have for any $u \in (0, \omega)$

$$\begin{aligned} &\boldsymbol{P}\{Y(1 - Z_{a,b})^{1/\tau} > u\} \\ &= \int_0^\omega [1 - B((u/s)^\tau, b, a)]\, dH(s) \\ &= \overline{H}(u) \int_0^{w(u)[\omega-u]} [1 - B([1 + s/(uw(u))]^{-\tau}, b, a)]\, dH_u(s). \end{aligned}$$

The assumption $H \in MDA(\Lambda)$ implies

$$\lim_{u\uparrow\omega}[H_u(t) - H_u(s)] = \exp(-s) - \exp(-t), \quad \forall s, t \in \mathbb{R}, s \geq t,$$

and

$$\lim_{u\uparrow\omega} w(u)[\omega - u] = \infty, \quad \lim_{u\uparrow\omega} uw(u) = \infty.$$

Consequently

$$\lim_{u\uparrow\omega}\left(\frac{uw(u)}{\tau}\right)^a \Big[1 - B([1 + s/(uw(u))]^{-\tau}, b, a)\Big] = \frac{\Gamma(a+b)}{a\Gamma(a)\Gamma(b)} s^a, \quad \forall s \in (0, \infty),$$

hence we obtain further

$$\liminf_{u\uparrow\omega}\left(\frac{uw(u)}{\tau}\right)^a \boldsymbol{P}\{Y(1 - Z_{a,b})^{1/\tau} > u\} \geq \frac{\Gamma(a+b)}{a\Gamma(a)\Gamma(b)} \int_{-\infty}^{\infty} s^a\, d(\exp(-s)).$$

The same upper bound can be shown for the lim sup of the left hand side above using Lemma 4.3 of Hashorva (2006).
ii) Breiman's Lemma implies as $u \to \infty$

$$\begin{aligned}\boldsymbol{P}\{Y(1-Z_{a,b})^{1/\tau} > u\} &= (1+o(1))\overline{H}(u)\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\int_0^1 x^{\alpha/\tau}x^{b-1}(1-x)^{a-1}\,dx \\ &= (1+o(1))\overline{H}(u)\frac{\Gamma(a+b)\Gamma(b+\alpha/\tau)}{\Gamma(b)\Gamma(a+b+\alpha/\tau)}.\end{aligned}$$

iii) Denote again by $B(y,a,b), y \in (0,1)$ the distribution function of $Z_{a,b}$ and put

$$B_u(y) := \tau_u^{-a} B(y\tau_u, a, b), \quad y \in [0,1], \quad \tau_u := \tau(1-u).$$

We may write for any $u \in (0,1)$

$$\boldsymbol{P}\{Y(1-Z_{a,b})^{1/\tau} > u\} = \tau_u^a \overline{H}(u)\int_0^{1-u^\tau} \frac{\overline{H}(u(1-y\tau_u)^{-1/\tau})}{\overline{H}(u)}\,dB_u(y).$$

Since for any $y \in (0,1)$

$$\lim_{u\uparrow 1}\frac{\overline{H}(u(1-y\tau_u)^{-1/\tau})}{\overline{H}(u)} = (1-y)^\alpha, \quad \text{and} \lim_{u\uparrow 1}\tau_u^{-a}B(y\tau_u,a,b) = \frac{y^a\Gamma(a+b)}{a\Gamma(a)\Gamma(b)}$$

applying Lemma 4.2 of Hashorva (2006) we obtain

$$\lim_{u\uparrow 1}\frac{\boldsymbol{P}\{Y(1-Z_{a,b})^{1/\tau} > u\}}{\tau_u^a[1-H(u)]} = \frac{\Gamma(\alpha+1)\Gamma(a+b)}{\Gamma(b)\Gamma(\alpha+a+1)},$$

hence the proof is complete. □

Proof of Theorem 1 i) Let $V \sim Beta((d-1)/p, 1/p)$ be independent of the associated random radius R. Using (2) we have for any $u > 0$

$$\boldsymbol{P}\{X_i > u\} = \frac{1}{2}\boldsymbol{P}\{R(1-V)^{1/p} > u\}, \quad i = 1,\ldots,d. \tag{30}$$

In view of Theorem 8 we obtain taking $\tau = p$

$$\lim_{u\uparrow\omega}\frac{1-G_i(u+x/w(u))}{1-G_i(u)} = \lim_{u\uparrow\omega}\frac{1-F(u+x/w(u))}{1-F(u)} = \exp(-x), \quad \forall x \in \mathbb{R},$$

hence $G_i, 1 \le i \le d$ is in the Gumbel max-domain of attraction with the scaling function w. Since $F \in MDA(\Lambda)$ implies (5), then (4) follows. Consequently the components of the sample maxima are asymptotically independent, hence $G \in MDA(Q)$ with Q a distribution function on $\mathbb{R}^d$ with independent unit Gumbel marginal distributions.
ii) By (29) and (30) it follows that $G_i \in MDA(\alpha+(d-1)/p), 1 \le i \le d$. Since the upper endpoint of F is finite we have that G has all marginal distributions with finite upper endpoint ω. In view of (2) the components of $\boldsymbol{X}$, say $X_i, X_k, i \ne k$ cannot be both extreme (near enough to ω) with a non-zero probability, implying $\kappa(X_1, X_2) = 0$. Hence, the sample maxima has asymptotic independent components. □

Proof of Theorem 2 The assumptions imply that

$$\boldsymbol{X} \stackrel{d}{=} R\boldsymbol{U}_d \stackrel{d}{=} R_{d,p}\boldsymbol{U}_d,$$

with $R_{d,p}$ independent of $\boldsymbol{U}_d$.
i) $\Rightarrow$ *ii)* In view of (2) we have

$$R_{i,p} \stackrel{d}{=} R_{d,p}(1-V_i)^{1/p},$$

with $V_i \sim Beta((d-i)/p, i/p)$ being further independent of the random radius $R_{d,p}$. Applying Lemma 8 we obtain taking $\tau = p$

$$\boldsymbol{P}\{R_{i,p} > u\} \quad = \quad (1+o(1))\frac{\Gamma(d/p)\Gamma((i+\alpha)/p)}{\Gamma(i/p)\Gamma((d+\alpha)/p)}\boldsymbol{P}\{R_{d,p} > u\}, \quad u \to \infty.$$

In view of (30) X_1 is regularly varying with index $\alpha > 0$ is equivalent with $R_{1,p}$ is regularly varying with the same index α, hence the claim follows.
iii) $\Rightarrow$ *ii)* Using the fact that $|X_1|^p = R_{1,p}^p$ and X_1 is symmetric about 0 establishes the proof.
iii) $\Rightarrow$ *i)* Clearly *iii)* includes *i)*.
iv) $\Rightarrow$ *v)* $\Rightarrow$ *ii)* This follows easily by the definition of the regular variation.
ii) $\Rightarrow$ *i)* Let $\boldsymbol{Z} = (Z_1, \ldots, Z_d)^\top$ be a L_p-norm spherical random vector as in Example 1 being further independent of $\boldsymbol{X}$ and let

$$\widetilde{R}_{i,p} := \Big(\sum_{j=1}^{i}|Z_i|^p\Big)^{1/p}, \quad 1 \le i \le d$$

be the i-th random radius associated with $\boldsymbol{Z}$. By the assumptions it follows easily that $R_{1,p}^p$ is regularly varying with index α/p. Since

$$\widetilde{R}_{i,p}^p \sim Gamma(i/p, 1/p), \quad 1 \le i \le p$$

and $\widetilde{R}_{i,p}$ is independent of $R_{1,p}$ Lemma 7 implies that the product $(\widetilde{R}_{d,p}R_{1,p})^p$ is regular varying with positive index α/p.
Let $V \sim Beta(1/p, (d-1)/p)$ be independent of $\boldsymbol{Z}$ and $\boldsymbol{X}$. Now, the stochastic representation (2) implies

$$(\widetilde{R}_{d,p}R_{1,p})^p \stackrel{d}{=} \widetilde{R}_{d,p}^p(R^pV) \stackrel{d}{=} \widetilde{R}_{d,p}^pVR_{d,p}^p \stackrel{d}{=} \widetilde{R}_{1,p}^pR_{d,p}^p,$$

consequently $\widetilde{R}_{1,p}^pR_{d,p}^p$ is regularly varying with index α/p.
We have $\widetilde{R}_{1,p}^p \sim Gamma(1/p, 1/p)$ with $\widetilde{R}_{1,p}^p$ independent of $R_{d,p}^p = R^p$. Applying again Lemma 7 we have that $R_{d,p}^p$ is regularly varying with positive parameter α/p, thus the proof follows. □

Proof of Theorem 4 i) Let i, j be fixed with $1 \le i < j \le d$. By the definition both X_i and X_j have distribution function with the same upper endpoint ω. Consequently, if $\omega < \infty$ then X_i and X_j cannot be close to ω with non-zero probability, i.e., $\boldsymbol{P}\{X_i > \omega - \varepsilon, X_j > \omega - \varepsilon\} = 0$ for some $\varepsilon > 0$ small enough, hence $\kappa(X_i, X_j) = 0$ for this case.
ii) If $\omega = \infty$ and $1 - F$ is rapidly varying, then in view of condition (20) the associated random radius $R_{\{i,j\},p} := (|X_i|^p + |X_j|^p)^{1/p}$ has a rapidly varying distribution function. Hence, the proofs follows then using (3) and (20). □

Proof of Theorem 5 i) Since the distribution function F is in the max-domain of attraction of Λ and (20) is supposed to hold, then Theorem 8 implies that $X_i, 1 \le i \le d$ has distribution function in the same max-domain of attraction with the scaling function w. If $\omega < \infty$ using further Theorem 4 we get that the sample maxima has asymptotic independent components. If $\omega = \infty$ and $\lambda_i = \lambda >$

$0, 1 \leq i \leq d$ then Theorem 8 implies the distribution functions of $\boldsymbol{X}_i, 1 \leq i \leq d$ have the same asymptotic tail behaviour (up to some constant), hence the sample maxima has asymptotic independent components, thus the proof follows easily.
ii) Again using Theorem 8 we have that $X_i, 1 \leq i \leq d$ has distribution function in the max-domain of attraction of $\Psi_{\alpha+\lambda_i}$. Since $\omega < \infty$ the components then X_i, X_j cannot be both near to ω for any $1 \leq i < j \leq d$, implying that the sample maxima has independent components, thus the proof is complete. □

PROOF OF THEOREM 6 The proof is similar to the proof of Theorem 2 using further Theorem 8. □

Acknowledgement: I would like to thank Professor Samuel Kotz for sending many helpful manuscripts, useful comments, corrections and new ideas. Many thanks are due to Dr. Marco Collenberg for some corrections and confirming the proof of Lemma 6.1. Several important references were kindly provided by Dr. Simon Rentzmann and Professor Szabłowski.

REFERENCES

[1] Basrak, B., Davis, R.A., and Mikosch, T. (2002) A Characterization of multivariate regular variation. *Ann. Appl. Prob.* **12**,3, 908–920.
[2] Berman, M.S. (1982) Sojourns and extremes of stationary processes. *Ann. Probability* **10**, 1–46.
[3] Berman, M.S. (1992) *Sojourns and Extremes of Stochastic Processes.* Wadsworth & Brooks/ Cole.
[4] Breiman, L. (1965) On some limit theorems similar to arc-sin law. *Theory of Probab. Appl.* **10**, 323–331.
[5] Cambanis, S., Huang, S., and Simons, G. (1981) On the theory of elliptically contoured distributions. *J. Multivariate Anal.* **11**, 368–385.
[6] Carnal, H. (1970) Die konvexe Hülle von n rotations-symmetrisch verteilten Punkten. *Z. Wahrscheinlichkeitstheorie Verw. Geb.* **15**, 168–176.
[7] de Haan, L. (1970) *On Regular Variation and its Applications to the Weak Convergence of Sample Extremes.* Mathematisch Centrum Amsterdam, The Netherlands.
[8] Eddy, W.F., and Gale, J.D. (1981) The convex hull of a spherically symmetric sample. *Advances in Applied Probability*, **13**, 751–763.
[9] Falk, M., Hüsler, J., and Reiss R.-D. (2004) *Laws of Small Numbers: Extremes and Rare Events.* DMV Seminar **23**, 2-nd edition, Birkhäuser, Basel.
[10] Fang, K.-T., and Fang, Bi-Qi. (1990) Generalised symmetrized Dirichlet distributions. In Statistical Inference in Elliptically Contoured and Related Distributions, K.T. Fang and T.W. Anderson, eds, Allerton Press, New York, pp. 127–136.
[11] Fang, K.-T., Kotz, S., and Ng, K.-W. (1990) *Symmetric Multivariate and Related Distributions.* Chapman and Hall, London.
[12] Feller, W. (1966) *An Introduction to Probability Theory and its Applications II.* Wiley, New York.
[13] Galambos, J. (1987) *The Asymptotic Theory of Extreme Order Statistics*, 2-nd edn. Malabar: Krieger.
[14] Gale, J.D. (1980) The Asymptotic Distribution of the Convex Hull of a Random Sample. *Ph.D. Thesis, Carnegie-Mellon University.*
[15] Gupta, A.K., and Song, D. (1997) L_p-norm spherical distributions, *J. Statist. Plann. Inference*, 60, 241–260.
[16] Hashorva, E. (2005) Extremes of asymptotically spherical and elliptical random vectors. *Insurance: Mathematics and Economics*, **36**, 3, 285–302.
[17] Hashorva, E. (2006) On the regular variation of elliptical random vectors. *Stat. Probab. Lett.* **76**,(14), 1427–1434.
[18] Hashorva, E. (2007a) Conditional limiting distribution of type III elliptical random vectors. *J. Multivariate Analysis.* **98**, 282–194.

[19] Hashorva, E. (2007b) Extremes and asymptotic dependence of elliptical random vectors. In: *Extreme Value Distributions.* Eds. Ahsanulah, M., and Kirmani, S. Nova Science Publishers. pp. 159–179.
[20] Hashorva, E., Kotz, S., and Kume, A. (2007) L_p-norm generalised symmetrised Sirichlet distributions. *Albanian Journal of Mathematics.* **1**, 31–56.
[21] Kotz, S. (1975) Multivariate distributions at a cross-road. In: *Statistical Distributions in Scientific Work* 1, G.P. Patil, S. Kotz, and J.K. Ord. D. Riedel Publishing Company eds., Dortrecht, 240–247.
[22] Kotz, S. (2004) Extremal Elliptical Distributions. *Manuscript.*
[23] Kotz, S., and Nadarajah, S. (2005) *Extreme Value Distributions, Theory and Applications.* Imperial College Press, London, United Kingdom. (Second Pringing).
[24] Leadbetter, M.R., Lindgren, G., and Rootzén, H. (1983) *Extremes and related properties of random sequences and processes.* Springer-Verlag, New York.
[25] Mikosch, T. (2005) How to model multivariate extremes if one must. *Statistica Neerlandica*, **59**, 324–338.
[26] Reiss, R-D. (1989) *Approximate Distributions of Order Statistics: With Applications to Non-parametric Statistics.* Springer, New York.
[27] Resnick, S.I. (1987) *Extreme Values, Regular Variation and Point Processes.* Springer, New York.
[28] Szabłowski, P.L. (1998) Uniform distributions on spheres in finite dimensional L_α and their generalizations. *J. Multivariate Anal.* **64**, 103–117.

University of Bern, Institute of Mathematical Statistics and Actuarial Sciences, Sidlerstrasse 5, CH-3012 Bern, Switzerland, and, Allianz Suisse Insurance Company, Laupenstrasse 27, CH-3001 Bern, Switzerland
E-mail address: enkelejd.hashorva@stat.unibe.ch
E-mail address: enkelejd.hashorva@Allianz-Suisse.ch

ALBANIAN JOURNAL
OF MATHEMATICS
Volume 1, Number 4, Pages 175–177
ISSN 1930-1235: (2007)

SPECIAL ISSUE ON ALGEBRA AND COMPUTATIONAL ALGEBRAIC GEOMETRY

A. ELEZI AND T. SHASKA

Algebraic geometry is one of the main branches of modern mathematics with roots from classical Italian geometers. Its modern flavor started with Grothendieck and continued with many illustrious algebraic geometers of the second half of the 20-th century. During the last twenty years, the subject has changed drastically due to developments of new computational techniques and access to better computing power. Such changes have spurred a new direction of algebraic geometry, the so called *computational algebraic geometry*. While there is no universal agreement among mathematicians that what exactly is computational algebraic geometry, loosely stated it includes the areas of algebraic geometry where computer algebra can be used to obtain explicit results. It is obvious that such area will be of deep impact and importance in the future mathematics. Furthermore, such new developments have made possible applications of algebraic geometry in areas such as coding theory, computer security and cryptography, computer vision, mathematical biology, and many more.

This special issue contains papers that cover classical mathematical problems from a computational viewpoint and problems in newer developments. We intentionally did not limit the papers in a narrow area. Instead, we tried to present a wide variety of topics. This special issue consists of ten papers on the following topics:

The paper by D. Haran and M. Jarden studies the embedding problem. Let K be an ample field, G a finite group, and L a finite Galois extension of K such that $Gal(L/K)$ is isomorphic to a subgroup of G. They prove that $K(x)$ has a Galois extension F which is regular over L such that $Gal(F/K(x)) \cong G$ and F has a K-place ϕ such that $\phi(x) \in K$ and $\phi(F) = L \cup \{\infty\}$.

The paper by M. Joswig, B. Sturmfels, and J. Yu explores the relationship between convexity in tropical geometry and notions of convexity in the theory of affine buildings, from a combinatorial and computational perspective. Results include a convex hull algorithm for the Bruhat–Tits building of $SL_d(K)$ and techniques for computing with apartments and membranes.

The paper by J. Hakim is on discrete series representations of p-adic groups associated to symmetric spaces. The purpose of this paper is study the natural symmetric space analogues of various notions related to discrete series representations of a p-adic group such as Schur's orthogonality relations and formal degrees.

A. Elezi in his paper focuses on toric fibrations and mirror symmetry. The relation between the quantum $\mathcal{D}$-modules of a smooth variety X and a toric bundle is studied. The author describes the relation completely when X is a semi-ample

complete intersection in a toric variety. In this case, all the relations in the small quantum cohomology ring of the bundle are obtained.

Q. Gashi studies toric varieties associated with root systems. Consider a root system R and the corresponding toric variety V_R whose fan is the Weyl fan and whose lattice of characters is given by the root lattice for R. The author proves the vanishing of the higher cohomology groups for certain line bundles on V_R by proving a purely combinatorial result for root systems. These results are related to a converse to Mazur's Inequality for split reductive groups.

In their paper A. Elkin and R. Pries show there exists a hyperelliptic curve of genus $g \geq 3$ with p-rank $g-3$ and a-number 1 in characteristic p when $p = 3$ or $p = 5$. The method of proof is to show that a generic point of the moduli space of hyperelliptic curves of genus 3 and p-rank 0 has a-number 1. When $p = 3$, it is also shown that this moduli space is irreducible.

There are two papers on theta functions of algebraic curves in this special issue. The first paper by E. Previato, T. Shaska, and S. Wijesiri studies relations among the classical thetanulls of cyclic curves, namely curves $\mathcal{X}$ (of genus $g(\mathcal{X}) > 1$) with an automorphism σ such that σ generates a normal subgroup of the group G of automorphisms, and $g(\mathcal{X}/\langle\sigma\rangle) = 0$. Relations between thetanulls and branch points of the projection are the object of much classical work, especially for hyperelliptic curves, and of recent work, in the cyclic case. In this paper the authors determine the curves of genus 2 and 3 in the locus $\mathcal{M}_g(G, \mathbf{C})$ for all G that have a normal subgroup $\langle\sigma\rangle$ as above, and all possible signatures $\mathbf{C}$, via relations among their thetanulls.

In the second paper about theta functions, Y. Kopeliovich describes modular equations of order prime p and theta functions. Let p be a prime integer and $\mathbf{H_g}$ be a collection of complex $g \times g$ matrices τ such that: i) $\tau = \tau^t$ i.e. τ is symmetric and ii) $Im\tau$ is a positive definite quadratic form.

Denote by $p\tau$ the multiplication of τ by p. In this paper it is described an explicit process to obtain algebraic identities between theta functions with integral characteristics evaluated at τ and $p\tau$. For $g = 1$ this produces modular equations between $\lambda(\tau), \lambda(p\tau)$ where $\lambda(\tau)$ is the invariant associated with elliptic curve generated by τ, described by the equation: $y^2 = x(x-1)(x-\lambda(\tau_1))$. Consequently, if $g > 1$ The algebraic identities we obtain might serve as a higher dimensional generalization for the one dimensional modular equations.

V. Ustimenko focuses on the extremal regular directed graphs without commutative diagrams and their applications in coding theory and cryptography. There are methods from computational algebra via Groebner basis which can be applied in the study of such graphs, even though they are not fully explored in this paper.

In the last paper, T. Shaska re-visits an old list of open problems in the area of computer algebra and algebraic curves. That list first appeared in 2003, in the ACM, *SIGSAM Bulletin*, *Comm. Comp. Alg.*. It was a list of problems on algebraic curves which could be approached computationally. Some of those problems were solved and many papers from various authors were written based on that modest paper. The author has updated the list with new problems and has included some problems on higher dimensional varieties.

Acknowledgements: We sincerely thank all the authors for their contributions of this special issue. Such contributions provide a wide view of computational

algebraic geometry. We also thank the anonymous referees for all their work going through all the papers.

Artur Elezi
American University

Tanush Shaska
Oakland University

ALBANIAN JOURNAL
OF MATHEMATICS
Volume 1, Number 4, Pages 179–185
ISSN 1930-1235: (2007)

REGULAR LIFTING OF COVERS OVER AMPLE FIELDS

D. HARAN AND M. JARDEN

Abstract. Let K be an ample field, G a finite group, and L a finite Galois extension of K such that $\mathrm{Gal}(L/K)$ is isomorphic to a subgroup of G. We prove that $K(x)$ has a Galois extension F which is regular over L such that $\mathrm{Gal}(F/K(x)) \cong G$ and F has a K-place φ such that $\varphi(x) \in K$ and $\varphi(F) = L \cup \{\infty\}$.

1. Introduction

Colliot-Thélène [CoT00] uses the technique of Kollár, Miyaoka, and Mori to prove the following result.

Theorem A: *Let K be an ample field of characteristic 0, x a transcendental element over K, and G a finite group. Then there is a Galois extension F of $K(x)$ with Galois group G, regular over K.*

Here K is said to be **ample** if every absolutely irreducible curve defined over K with a K-rational simple point has infinitely many K-rational simple points.

In fact, Colliot-Thélène proves a stronger version, still under the assumption that K is ample and $\mathrm{char}(K) = 0$:

Theorem B: *Given a Galois extension L/K with Galois group Γ which is a subgroup of G, there exist a Galois extension F of $K(x)$ with $\mathrm{Gal}(F/K(x)) \cong G$ and a place φ that fixes the elements of K and the residue field extension of $F/K(x)$ under φ is L/K.*

Case $\Gamma = G$ of Theorem B means that K has the arithmetic lifting property of Beckmann and Black [Bla99].

Since the results of Kollár, Miyaoka, and Mori are valid only in characteristic 0, Colliot-Thélène's proof works only in this case. Nonetheless, Theorem A holds in arbitrary characteristic ([Har87, Corollary 2.4] for complete fields, [Pop96, Main Theorem A]; see also [Liu95] and [HaV96]). Theorem B can be deduced for arbitrary characteristic from Théorème 1.1 of [MoB01]. The proof of that paper uses methods of formal patching.

Here we use algebraic patching to prove Theorem B for arbitrary characteristic. In fact, the main ingredient of the proof is almost contained in [HaJ98]. Therefore

Both authors were partially supported by the Minkowski Center for Geometry at Tel Aviv University and the Mathematical Sciences Research Institute, Berkeley.

this note can be considered a sequel to [HaJ98]; a large portion of it recalls the situation and facts considered there.

The idea (displayed in our Lemma 3.1) to use the embedding problem $G \ltimes G \to G$ in order to obtain the arithmetic lifting property has been used in [Pop99]; we are grateful to F. Pop for making his note available to us.

2. Embedding problems and decomposition groups

Let K/K_0 be a finite Galois extension with Galois group Γ. Let x be a transcendental element over K. Put $E_0 = K_0(x)$. Suppose that Γ acts (from the right) on a finite group G; let $\Gamma \ltimes G$ be the corresponding semidirect product and $\pi\colon \Gamma \ltimes G \to \Gamma$ the canonical projection. We call

$$\pi\colon \Gamma \ltimes G \to \Gamma = \mathrm{Gal}(K/K_0) \tag{1}$$

a **finite constant split embedding problem**. A **solution** of (1) is a Galois extension F of E_0 such that $K \subseteq F$, $\mathrm{Gal}(F/E_0) = \Gamma \ltimes G$, and π is the restriction map $\mathrm{res}_K\colon \mathrm{Gal}(F/E_0) \to \mathrm{Gal}(K/K_0)$.

In [HaJ98, Theorem 6.4] we reprove the following result of F. Pop [Pop96]:

Proposition 2.1. *Let K_0 be an ample field. Then each finite constant split embedding problem (1) has a solution F such that F has a K-rational place φ such that $\varphi(x) \in K_0 \cup \{\infty\}$ (in particular, F/K is regular).*

In this section we show that the proof of Proposition 5.2 in [HaJ98] yields a stronger assertion.

We denote the residue field of a place φ of a field F by $\bar{F}_\varphi$.

Lemma 2.2. *Let F be a solution of (1). Put $F_0 = F^\Gamma$. Let $\varphi\colon F \to \widetilde{K_0} \cup \{\infty\}$ be a K-place with $\varphi(x) \in K_0 \cup \{\infty\}$. Assume that φ is unramified in F/E_0 and let D_φ be its decomposition group in F/E_0. Then $K \subseteq \bar{F}_\varphi$ and the following assertions are equivalent:*

(a) $K = \bar{F}_\varphi$ and $\Gamma = D_\varphi$;
(b) $D_\varphi \subseteq \Gamma$;
(c) $K_0 = \bar{F}_{0,\varphi}$;
(d) $K = \bar{F}_\varphi$ and $\varphi(f^\gamma) = \varphi(f)^\gamma$ for each $\gamma \in \Gamma$ and $f \in F$ with $\varphi(f) \neq \infty$.

Proof. Since $K \subseteq F$, we have $K = \bar{K}_\varphi \subset \bar{F}_\varphi$. Since the inertia group of φ in F/E_0 is trivial, we have an isomorphism $\theta\colon D_\varphi \to \mathrm{Gal}(\bar{F}_\varphi/K_0)$ given by

$$\varphi(f^\gamma) = \varphi(f)^{\theta(\gamma)}, \qquad \gamma \in D_\varphi,\ f \in F,\ \varphi(f) \neq \infty. \tag{2}$$

Hence, $|D_\varphi| = [\bar{F}_\varphi : K_0] \geq [K : K_0] = |\Gamma|$. This gives (a) $\Leftrightarrow$ (b).

Since φ is unramified over E_0, the decomposition field F^{D_φ} is the largest intermediate field of F/E_0 mapped by φ into $K_0 \cup \{\infty\}$, and hence (b) $\Leftrightarrow$ (c).

Clearly (d) $\Rightarrow$ (c). If $\bar{F}_\varphi = K$, then $f^\gamma = \varphi(f^\gamma) = \varphi(f)^{\theta(\gamma)} = f^{\theta(\gamma)}$ for all $f \in K$ and $\gamma \in D_\varphi$ (by (2)). Hence, $\theta(\gamma) = \gamma$ for all $\gamma \in D_\varphi$. Applying (2) once more, we have $\varphi(f^\gamma) = \varphi(f)^{\theta(\gamma)} = \varphi(f)^\gamma$ for each $f \in F$ with $\varphi(f) \neq \infty$ and $\gamma \in D_\varphi$. Consequently, (a) $\Rightarrow$ (d).

□

Remark 2.3. *Let K_0 be an ample field and F a solution of (1). Suppose F has a K-rational place φ unramified over E_0 such that $\varphi(x) \in K_0 \cup \{\infty\}$ and Γ is the decomposition group of φ in F/E_0. Then F has infinitely many such places.*

Proof. Indeed, put $F_0 = F^\Gamma$. Recall that F_0 is regular over K_0. By Lemma 2.2,

(a) the assumption is that there is a K_0-place $\varphi\colon F_0 \to K_0$ unramified over $K_0(x)$, and

(b) we have to show that there are infinitely many such places.

But (a) $\Rightarrow$ (b) is a property of an ample field.

□

Proposition 2.4. *Let K_0 be an ample field. Then each finite constant split embedding problem (1) has a solution F with a K-rational place φ of F unramified over E_0 such that $\varphi(x) \in K_0 \cup \{\infty\}$ and Γ is the decomposition group of φ in F/E_0.*

Proof. Put $E = K(x) = KK_0(x)$.

Part A: As in the proof of [HaJ98, Theorem 6.4], we first assume that *K_0 is complete with respect to a non-trivial discrete ultrametric absolute value $|\ |$, with infinite residue field and K/K_0 is unramified.*

In this case [HaJ98, Proposition 5.2] proves Proposition 2.1. Claim C of that proof shows that, for every $b \in K_0$ with $|b| > 1$, $x \to b$ extends to a K-homomorphism $\varphi_b\colon R \to K$, where R is the principal ideal ring $K\{\frac{1}{x-c_i} \mid i \in I\}$ and the c_i's are properly chosen elements of K. From there it extends to a K-place $\varphi_b\colon Q \to K \cup \{\infty\}$ of the $Q = \mathrm{Quot}(R)$. Furthermore, [HaJ98, Lemma 1.3(b)] gives an E-embedding $\lambda\colon F \to Q$. The compositum $\varphi = \varphi_b \circ \lambda$ is a K-rational place of F. Excluding finitely many b's we may assume that φ is unramified over E_0. To verify that φ satisfies condition (d) of Lemma 2.2, we first recall the relevant facts from [HaJ98].

(a) [HaJ98, Proposition 5.2, Construction B] The group $\Gamma = \mathrm{Gal}(K/K_0)$ lifts isomorphically to $\mathrm{Gal}(E/E_0)$. By the choice of the c_i we have $\big(\frac{1}{x-c_i}\big)^\gamma = \frac{1}{x-c_i^\gamma}$, for each $\gamma \in \Gamma$. It follows that Γ continuously acts on R in the following way

$$\Big(a_0 + \sum_{i\in I}\sum_{n=1}^{\infty} a_{in}\big(\frac{1}{x-c_i}\big)^n\Big)^\gamma = a_0^\gamma + \sum_{i\in I}\sum_{n=1}^{\infty} a_{in}^\gamma\big(\frac{1}{x-c_i^\gamma}\big)^n.$$

This action induces an action of Γ on Q.

(b) [HaJ98, (7) on p. 334] The above mentioned action of Γ on Q defines an action of Γ on the Q-algebra

$$N = \mathrm{Ind}_1^G Q = \Big\{ \sum_{\theta\in G} a_\theta\theta \mid\ a_\theta \in Q\Big\}$$

in the following way:

$$\Big(\sum_{\theta\in G} a_\theta\theta\Big)^\gamma = \sum_{\theta\in G} a_\theta^\gamma\theta^\gamma \qquad a_\theta \in Q,\ \gamma \in \Gamma.$$

Furthermore, the field F is a subring of N [HaJ98, p. 332] and Γ acts on it by restriction from N [HaJ98, Proof of Proposition 1.5, Part A].

(c) The embedding $\lambda\colon F \to Q$ is the restriction to F of the projection

$$\sum_{\theta\in G} a_\theta\theta \mapsto a_1$$

from $N = \mathrm{Ind}_1^G Q$ onto Q [HaV96, Proposition 3.4].

(d) The place $\varphi_b\colon Q \to K \cup \{\infty\}$ is induced from the evaluation homomorphism $\varphi_b\colon R \to K$ given by [HaJ98, Remark 3.5]

$$\varphi_b\Big(a_0 + \sum_{i\in I}\sum_{n=1}^{\infty} a_{in}\big(\frac{1}{x-c_i}\big)^n\Big) = a_0 + \sum_{i\in I}\sum_{n=1}^{\infty} a_{in}\big(\frac{1}{b-c_i}\big)^n.$$

In order to prove condition (d) of Lemma 2.2 it suffices to show that both λ and φ_b are Γ-equivariant.

Let $f = \sum_{\theta\in G} a_\theta\theta \in F \subseteq N$. Then, by (b) and (c),

$$\lambda(f^\gamma) = \lambda\Big(\sum_{\theta\in G} a_\theta^\gamma\theta^\gamma\Big) = a_1^\gamma = \Big(\lambda\Big(\sum_{\theta\in G} a_\theta\theta\Big)\Big)^\gamma = \lambda(f)^\gamma.$$

Furthermore, let $r = a_0 + \sum_{i\in I}\sum_{n=1}^{\infty} a_{in}\big(\frac{1}{x-c_i}\big)^n \in R$. By (a) and (d),

$$\begin{aligned}\varphi_b(r^\gamma) &= \varphi_b\Big(a_0^\gamma + \sum_{i\in I}\sum_{n=1}^{\infty} a_{in}^\gamma\big(\frac{1}{x-c_i^\gamma}\big)^n\Big) = a_0^\gamma + \sum_{i\in I}\sum_{n=1}^{\infty} a_{in}^\gamma\big(\frac{1}{b-c_i^\gamma}\big)^n\\ &= \Big(a_0 + \sum_{i\in I}\sum_{n=1}^{\infty} a_{in}\big(\frac{1}{b-c_i}\big)^n\Big)^\gamma = \varphi_b(r)^\gamma.\end{aligned}$$

Thus φ_b is Γ-equivariant.

Part B: *K_0 is an arbitrary ample field.* As in the proof of [HaJ98, Theorem 6.4] let $\hat{K}_0 = K_0((t))$ be the field of formal power series in t over K_0. Then $\hat{K} = K\hat{K}_0$ is an unramified extension of $\hat{K}_0$ with Galois group Γ and infinite residue field.

By Part A, $\hat{K}_0(x)$ has a Galois extension $\hat{F}$ which contains $\hat{K}(x)$, such that $\mathrm{Gal}(\hat{F}/\hat{K}_0(x)) = \Gamma \ltimes G$ and the restriction map $\mathrm{Gal}(\hat{F}/\hat{K}_0(x)) \to \mathrm{Gal}(K/K_0)$ is the projection $\pi\colon \Gamma\ltimes G \to \Gamma$. Furthermore, there is $b \in \hat{K}_0$ such that the place $x \to b$ of $\hat{K}_0(x)$ extends to an unramified $\hat{K}$-place $\hat{\varphi}\colon \hat{F} \to \hat{K} \cup \{\infty\}$ and $\hat{\varphi}(\hat{F}^\Gamma) = \hat{K}_0$. Put $m = |G|$.

Use the Weak Approximation to find $y \in \hat{F}^\Gamma$ mapped by the m distinct extensions of $x \to b$ to $\hat{F}^\Gamma$ into m distinct elements of the separable closure of $\hat{K}_0$; then $\hat{F}^\Gamma = \hat{K}_0(x, y)$.

Thus there exist polynomials $f \in \hat{K}_0[X, Z]$, $g \in \hat{K}_0[X, Y]$, elements $z \in \hat{F}$, $y \in \hat{F}^\Gamma$, and elements $b, c \in \hat{K}_0$, such that the following conditions hold:

(3a) $\hat{F} = \hat{K}_0(x, z)$, $f(x, Z) = \mathrm{irr}(z, \hat{K}_0(x))$; we identify $\mathrm{Gal}(f(x, Z), \hat{K}_0(x))$ with $\mathrm{Gal}(\hat{F}/\hat{K}_0(x))$;

(3b) $\hat{F}^\Gamma = \hat{K}_0(x, y)$, whence $\hat{F} = \hat{K}(x, y)$, and $g(x, Y) = \mathrm{irr}(y, \hat{K}_0(x))$; therefore $g(X, Y)$ is absolutely irreducible;

(3c) $\mathrm{discr}(g(b, Y)) \neq 0$ and $g(b, c) = 0$.

All of these objects depend on only finitely many parameters from $\hat{K}_0$. Hence, there are $u_1, \dots, u_n \in \hat{K}_0$ such that the following conditions hold:

(4a) $F = K_0(\mathbf{u}, x, z)$ is a Galois over $K_0(\mathbf{u}, x)$, the coefficients of $f(X, Z)$ lie in $K_0[\mathbf{u}]$, $f(x, Z) = \mathrm{irr}(z, K_0(\mathbf{u}, x))$, and $\mathrm{Gal}(f(x, Z), K_0(\mathbf{u}, x)) = \mathrm{Gal}(f(x, Z), \hat{K}_0(x))$;

(4b) the coefficients of g lie in $K[\mathbf{u}]$; hence $g(x, Y) = \mathrm{irr}(y, K_0(\mathbf{u}, x))$; furthermore, $K_0(\mathbf{u}, x, y) = F^\Gamma$;

(4c) $b, c \in K_0[\mathbf{u}]$, $\mathrm{discr}(g(b, Y)) \neq 0$, and $g(b, c) = 0$.

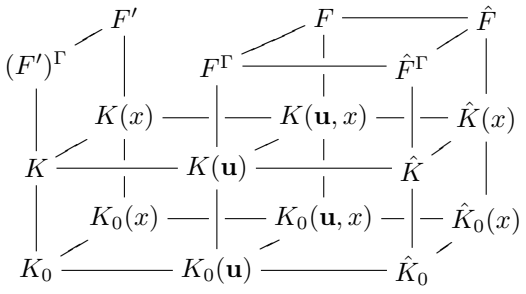
Since $\hat{K}_0$ has a K-rational place, namely, $x \to 0$, the field $\hat{K}_0$ and therefore also $K_0(\mathbf{u})$ are regular extensions of K_0. Thus, $\mathbf{u}$ generates an absolutely irreducible variety $U = \mathrm{Spec}(K_0[\mathbf{u}])$ defined over K_0. By Bertini-Noether [FrJ05, Proposition 9.4.3], the variety U has a nonempty Zariski open subset U' such that for each $\mathbf{u}' \in U'$ the K_0-specialization $\mathbf{u} \to \mathbf{u}'$ extends to a K-homomorphism $' \colon K[\mathbf{u}, x, z, y] \to K[\mathbf{u}', x, z', y']$ such that the following conditions hold:

(5a) $f'(x, z') = 0$, the discriminant of $f'(x, Z)$ is not zero, and $F' = K_0(\mathbf{u}', x, z')$ is the splitting field of $f'(x, Z)$ over $K_0(\mathbf{u}', x)$; in particular $F'/K_0(\mathbf{u}', x)$ is Galois;

(5b) $g'(X, Y)$ is absolutely irreducible and $g'(x, y') = 0$; so $g'(x, Y) = \mathrm{irr}(y', K(\mathbf{u}', x))$; furthermore, $K_0(\mathbf{u}', x, y') = (F')^\Gamma$;

(5c) $b', c' \in K_0[\mathbf{u}']$ and $\mathrm{discr}(g'(b', Y)) \neq 0$ and $g'(b', c') = 0$.

By assumption, K_0 is ample, so K_0 is existentially closed in $\hat{K}_0$ [Pop96, Prop. 1.1]. Since $\mathbf{u} \in U(\hat{K}_0)$, there is a $\mathbf{u}' \in U(K_0)$. Now repeat the end of the proof of [HaJ98, Lemma 6.2] (from "By (5a), the homomorphism...") to conclude that F' is a solution of (1).



Condition (5c) ensures that the place $x \to b'$ of $K_0(x)$ is unramified in $(F')^\Gamma$, hence in F', and extends to a K_0-rational place of $(F')^\Gamma$. This ends the proof by Lemma 2.2.

□

3. Lifting property over ample fields

Consider a subgroup Γ of a finite group G, let Γ act on G by the conjugation in G

$$g^\gamma = \gamma^{-1} g \gamma.$$

and consider the semidirect product $\Gamma \ltimes G$. To fix notation,

$$\Gamma \ltimes G = \{(\gamma, g) \mid \gamma \in \Gamma, g \in G\}$$

and the multiplication on $\Gamma \ltimes G$ is defined by

$$(\gamma_1, g_1)(\gamma_2, g_2) = (\gamma_1 \gamma_2, g_1^{\gamma_2} g_2).$$

Notice the isomorphism $\Gamma \ltimes G \cong \Gamma \times G$ given by $(\gamma, g) \mapsto (\gamma, \gamma g)$ and the epimorphism $\rho\colon \Gamma \ltimes G \to G$ given by $(\gamma, g) \mapsto \gamma g$. Let $N = \mathrm{Ker}(\rho)$.

Lemma 3.1. *Let K_0 be a field, K a Galois extension of K_0 with Galois group Γ, and x a transcendental element over K_0. Assume that (1) has a solution $\hat{F}$ with a K-rational place $\hat{\varphi}$ of $\hat{F}$ unramified over $K_0(x)$ such that $\hat{\varphi}(x) \in K_0 \cup \{\infty\}$ and Γ is the decomposition group of $\hat{\varphi}$ in $\hat{F}/K_0(x)$. Let $F = \hat{F}^N$ and let φ be the restriction of $\hat{\varphi}$ to F. Then*

(6a) F is a Galois extension of $K_0(x)$ and $\mathrm{Gal}(F/K_0(x)) \cong G$;

(6b) F/K_0 is a regular extension;

(6c) φ represents a prime divisor $\mathfrak{p}$ of F/K_0 with decomposition group Γ in $F/K_0(x)$ and residue field K.

Proof. By assumption, $\hat{F}$ is a Galois extension of $K_0(x)$ containing K, with Galois group $\Gamma \ltimes G$ such that the restriction $\mathrm{Gal}(\hat{F}/K_0(x)) \to \mathrm{Gal}(K/K_0)$ is the projection $\Gamma \ltimes G \to \Gamma$, and $\hat{F}/K$ is regular. Furthermore, $\hat{\varphi}\colon \hat{F} \to K$ is a K-place unramified over $K_0(x)$, with decomposition group $\Delta = \{(\gamma, 1) \mid \gamma \in \Gamma\} \cong \Gamma$ in $\hat{F}/K_0(x)$ and residue field extension K/K_0. In particular, $\hat{F}$ is regular over K.

From the definition of F we get (6a) and $\rho(\Delta) = \Gamma \leq G$ is the decomposition group of the restriction $\varphi\colon F \to K$ of $\hat{\varphi}$ to F. Since $|\Delta| = [K : K_0]$, the residue field of φ is K. Since $\Gamma \ltimes G = NG$, the fields $F = \hat{F}^N$ and $K(x) = \hat{F}^G$ are linearly disjoint over $K_0(x)$. In addition, $FK = \hat{F}$ and $\hat{F}/K$ is regular. Therefore, F is regular over K_0. □

Lemma 3.1 together with Proposition 2.4 and Remark 2.3 yield the following result:

Theorem 3.2. *Let K_0 be an ample field, G a finite group, Γ a subgroup, K a Galois extension of K_0 with Galois group Γ, and x a transcendental element over K_0. Then there is a field F that satisfies (6a), (6b) and*

(6d) there are infinitely many prime divisors $\mathfrak{p}$ of F/K_0 with decomposition group Γ in $F/K_0(x)$ and residue field K.

Remark 3.3. *In case of $\Gamma = G$, Theorem 3.2 says that an ample field K_0 has the so-called* **arithmetic lifting property** *of Beckmann-Black [Bla99].*

Remark 3.4. *In the special case where K is a PAC field, it possible to refine Theorem 3.2. In this case if F is an arbitrary Galois extension of $K(x)$ regular over K and L/K is a Galois extension with Galois group isomorphic to a subgroup of* $\mathrm{Gal}(F/K(x))$*, there exists a place φ of F such that the residue field extension of $F/K(x)$ under φ is L [Deb99, Remark 3.3]. This stronger property of PAC fields does not hold for an arbitrary ample field K [CoT00, Appendix].*

References

[BLa99] E.V. Black, *Deformations of dihedral 2-group extensions of fields,* Transactions of the AMS **351** (1999), 3229–3241.

[CoT00] J.-L. Colliot-Thélène, *Rational connectedness and Galois cover of the projective line,* Annals of Mathematics **151** (2000), 359-373.

[Deb99] P. Dèbes, *Galois Covers with Prescribed Fibers: The Beckmann-Black Problem,* Ann. Scuola Norm. Sup. Pisa **28** (1999), 273–286.

[FrJ05] M. D. Fried and M. Jarden, *Field Arithmetic, Second Edition, revised and enlarged by Moshe Jarden,* Ergebnisse der Mathematik (3) **11**, Springer, Heidelberg, 2005.

[Har87] . Harbater, *Galois coverings of the arithmetic line,* in: Lecture Notes in Mathematics **1240**, 165–195, Springer-Verlag 1987.

[HaJ98] D. Haran and M. Jarden, *Regular split embedding problems over complete valued fields,* Forum Mathematicum **10** (1998), 329–351 .

[HaV96] D. Haran and H. Völklein, *Galois groups over complete valued fields,* Israel Journal of Mathematics, **93** (1996), 9–27.

[Liu95] Q. Liu, *Tout groupe fini est un groupe de Galois sur $\mathbb{Q}_p(T)$, d'après Harbater,* Contemporary Mathematics **186** (1995), 261–265.

[MoB01] L. Moret-Bailly, *Construction de revêtements de courbes pointées,* Journal of Algebra **240** (2001), 505–534.

[Pop96] F. Pop, *Embedding problems over large fields,* Annals of Mathematics **144** (1996), 1–34.

[Pop99] F. Pop, *On the BB theory,* an unpublished note, 2 pages.

Dan Haran
School of Mathematics,
Tel Aviv University,
Ramat Aviv, Tel Aviv 69978,
Israel
Email: haran@math.tau.ac.il

Moshe Jarden
School of Mathematics,
Tel Aviv University,
Ramat Aviv, Tel Aviv 69978,
Israel
Email: jarden@math.tau.ac.il

ALBANIAN JOURNAL
OF MATHEMATICS
Volume 1, Number 4, Pages 187–211
ISSN 1930-1235: (2007)

AFFINE BUILDINGS AND TROPICAL CONVEXITY

MICHAEL JOSWIG, BERND STURMFELS, AND JOSEPHINE YU

Abstract. The notion of convexity in tropical geometry is closely related to notions of convexity in the theory of affine buildings. We explore this relationship from a combinatorial and computational perspective. Our results include a convex hull algorithm for the Bruhat–Tits building of $\mathrm{SL}_d(K)$ and techniques for computing with apartments and membranes. While the original inspiration was the work of Dress and Terhalle in phylogenetics, and of Faltings, Kapranov, Keel and Tevelev in algebraic geometry, our tropical algorithms will also be applicable to problems in other fields of mathematics.

1. Introduction

Buildings were initially introduced by Tits [24] to provide a common geometric framework for all simple Lie groups, including those of exceptional type. The later work of Bruhat and Tits [5] showed that buildings are fundamental in a much wider context, for instance, for applications in arithmetic algebraic geometry. Among the affine buildings, the key example is the *Bruhat–Tits building* $\mathcal{B}_d$ of the special linear group $\mathrm{SL}_d(K)$ over a field K with a discrete non-archimedean valuation. An active line of research explores compactifications of the building $\mathcal{B}_d$; for example, see Kapranov [16] and Werner [25, 26].

Our motivation to study affine buildings stems from the connection to biology which was proposed in Andreas Dress' 1998 ICM lecture *The tree of life and other affine buildings* [9]. Dress and Terhalle [8] introduced *valuated matroids* as a combinatorial approximation of the building $\mathcal{B}_d$, thereby generalizing the familiar one-dimensional picture of an infinite tree for $d = 2$. In Section 4 we shall see that valuated matroids are equivalent to the *matroid decompositions* of hypersimplices of Kapranov [16, Definition 1.2.17], to the *tropical linear spaces* of Speyer [22], and to the *membranes* of Keel and Tevelev [17]. The latter equivalence, shown in [17, Theorem 4.11], will be revisited in Theorem 18 below.

We start out in Section 2 with a brief introduction to the Bruhat–Tits building $\mathcal{B}_d$ and to the notion of convexity in $\mathcal{B}_d$ which appears in work of Faltings [10]. For sake of concreteness we take K to be the field $\mathbb{C}(\!(z)\!)$ of formal Laurent series with complex coefficients. Our discussion revolves around the algorithmic problem of computing the convex hull of a finite set of points in the building $\mathcal{B}_d$. Here each point is a lattice which is represented by an invertible $d \times d$-matrix with entries in $K = \mathbb{C}(\!(z)\!)$. Our solution to this problem involves identifying their convex hull in $\mathcal{B}_d$ with a certain *tropical polytope*.

Tropical convexity was introduced by Develin and Sturmfels [7]. Tropical polytopes are contractible polytopal complexes which are dual to the regular polyhedral subdivisions of the product of two simplices. A review of tropical convexity will be

given in Section 4, along with some new results, extending a formula of Ardila [3], which characterize the nearest point projection onto a tropical polytope. In Section 4, we introduce tropical linear spaces, we represent them as tropical polytopes, and we identify them with membranes in $\mathcal{B}_d$. This allows us in Section 5 to reduce convexity in $\mathcal{B}_d$ to tropical convexity. In addition to our convex hull algorithm, we also study the related problems of intersecting apartments or, more generally, membranes. We prove the following result:

Theorem 1. *The min-convex hull of a finite set of lattices in $\mathcal{B}_d$ coincides with the standard triangulation of a tropical polytope in a suitable membrane. The max-convex hull coincides with the image of a max-tropical polytope under the nearest point map onto a min-tropical linear space.*

This is stated more precisely in Proposition 22. New contributions made by this paper include the triangulation of tropical polytopes in Theorem 11, the formulas for projecting onto tropical linear spaces in Theorem 15, a combinatorial proof for the Keel-Tevelev bijection in Theorem 18, and, most important of all, the algorithms in Sections 5 and 6.

Acknowledgments: We are indebted to Ulrich Görtz for careful reading and requiring several clarifications concerning our exposition. Michael Joswig was partially supported by Deutsche Forschungsgemeinschaft (FOR 565 *Polyhedral Surfaces*). Bernd Sturmfels was partially supported by the National Science Foundation (DMS-0456960), and Josephine Yu was supported by a UC Berkeley Graduate Opportunity Fellowship and by the IMA in Minneapolis.

2. The Bruhat–Tits building of $\mathrm{SL}_d(K)$

We review basic definitions concerning Bruhat–Tits buildings, following the presentations in [17, 18]. The most relevant section in the monograph by Abramenko and Brown is [1, §6.9]. Let $R = \mathbb{C}[\![z]\!]$ be the ring of formal power series with complex coefficients. Its field of fractions is the field $K = \mathbb{C}(\!(z)\!)$ of formal Laurent series with complex coefficients. Taking the exponent of the lowest term of a power series defines a valuation $\mathrm{val} : K^* \to \mathbb{Z}$. Note that R is the subring of K consisting of all field elements c with $\mathrm{val}(c) \geq 0$. What follows is completely general and works for other fields with a non-archimedean discrete valuation, notably the p-adic numbers, but to keep matters most concrete we fix $K = \mathbb{C}(\!(z)\!)$. We extend the valuation to K by setting $\mathrm{val}(0) = \infty$. If M is a matrix over K then $\mathrm{val}(M)$ denotes the matrix over $\mathbb{Z} \cup \{\infty\}$ whose entries are the values of the entries of M.

The vector space K^d is a module over the ring R. A *lattice* in K^d is an R-submodule generated by d linearly independent vectors in K^d. Each lattice Λ is represented as the image of a matrix M with d rows and $\geq d$ columns, with entries in K, having rank d. Two lattices $\Lambda_1, \Lambda_2 \subset K^d$ are *equivalent* if $c\Lambda_1 = \Lambda_2$ for some $c \in K^*$. Two equivalence classes of lattices are called *adjacent* if there are representatives Λ_1 and Λ_2 such that $z\Lambda_2 \subset \Lambda_1 \subset \Lambda_2$.

The *Bruhat–Tits building* of $\mathrm{SL}_d(K)$ is the flag simplicial complex $\mathcal{B}_d$ whose vertices are the equivalence classes of lattices in K^d and whose edges are the adjacent pairs of lattices. Being a *flag simplicial complex* means that a finite set of vertices forms a simplex if and only if any two elements in that set form an edge. The link of any lattice Λ in $\mathcal{B}_d$ is isomorphic to the simplicial complex of all chains of

subspaces in $\mathbb{C}^d = \Lambda/z\Lambda$. Thus the simplicial complex $\mathcal{B}_d$ is pure of dimension $d-1$, but it is not locally finite, since the residue field is $\mathbb{C}$. Our objective is to identify finite subcomplexes with a nice combinatorial structure which is suitable for reducing computations in $\mathcal{B}_d$ to tropical geometry.

If Λ_1 and Λ_2 are lattices then their R-module sum $\Lambda_1 + \Lambda_2$ is generated as an R-module by the union of generators of Λ_1 and Λ_2. And, since R is a principal ideal domain every finitely generated torsion-free R-module is free, whence $\Lambda_1 + \Lambda_2$ is a lattice. Further, the intersection $\Lambda_1 \cap \Lambda_2$ is a also lattice by duality in Lemma 2. These two operations give rise to two different notions of convexity on the Bruhat–Tits building $\mathcal{B}_d$. We say that a set $\mathcal{M}$ of lattices in $\mathcal{B}_d$ is *max-convex* if the set of all representatives for lattices in $\mathcal{M}$ is closed under finite R-module sums. We call $\mathcal{M}$ *min-convex* if that set is closed under finite intersections. If $\mathcal{L}$ is any subset of $\mathcal{B}_d$ then its *max-convex hull* $\mathrm{maxconv}(\mathcal{L})$ is the set of all lattices Λ in K^d such that Λ is the R-module sum of finitely many lattices in $\mathcal{L}$. Similarly, the min-convex hull $\mathrm{minconv}(\mathcal{L})$ is the set of all lattices Λ in K^d such that Λ is the intersection of finitely many lattices in $\mathcal{L}$. These notions of convexity give rise to the following problem in computational algebra:

Computational Problem A. Let $M_1, \ldots, M_s$ be invertible $d \times d$-matrices with entries in $K = \mathbb{C}(\!(z)\!)$, representing lattices $\Lambda_i = \mathrm{image}_R(M_i)$ in K^d. Compute both the min-convex hull and the max-convex hull of the lattices $\Lambda_1, \ldots, \Lambda_s$ in the Bruhat–Tits building $\mathcal{B}_d$.

The duality functor $\mathrm{Hom}_R(\,\cdot\,, R)$ reduces a min-convex hull computation to a max-convex hull computation and vice versa. Given any lattice Λ, we write $\Lambda^* = \mathrm{Hom}_R(\Lambda, R)$ for the dual lattice. Any R-module homomorphism $\Lambda \to R$ extends uniquely to a K-vector space homomorphism $K^d \to K$. Hence the free R-module Λ^* can be considered as a lattice in the dual vector space $(K^d)^* = \mathrm{Hom}_K(K^d, K)$, consisting of those elements that send Λ into R. For any unit $c \in K^*$, we have $(c\Lambda)^* = \frac{1}{c}(\Lambda^*)$. Since duality is inclusion-reversing, i.e. $\Lambda_1 \subset \Lambda_2$ implies $\Lambda_2^* \subset \Lambda_1^*$, it respects equivalence of lattices and adjacency of vertices in the building $\mathcal{B}_d$. Moreover, duality switches sums and intersections:

Lemma 2. *For any two lattices Λ_1, Λ_2 in K^d, we have $(\Lambda_1 + \Lambda_2)^* = \Lambda_1^* \cap \Lambda_2^*$ in $(K^d)^*$.*

Proof. The inclusion "$\subseteq$" is given by restricting any ring homomorphism $\phi : \Lambda_1 + \Lambda_2 \to R$ to Λ_1 and to Λ_2, respectively. The reverse inclusion "$\supseteq$" is given by identifying $\phi \in \Lambda_1^* \cap \Lambda_2^*$ with the map $f_1 + f_2 \mapsto \phi(f_1) + \phi(f_2)$ where $f_i \in \Lambda_i$. □

It is known that both the max-convex hull and the min-convex hull of $\Lambda_1, \ldots, \Lambda_s$ are finite simplicial complexes of dimension $\leq d-1$. This finiteness result is attributed by Keel and Tevelev [17, Lemma 4.11] to Faltings' paper on matrix singularities [10, Lemma 3].

Our usage of the prefixes "min" and "max" for convexity in $\mathcal{B}_d$ is consistent with the alternative representation of the Bruhat–Tits building in terms of additive norms on K^d. An *additive norm* is a map $N : K^d \to \mathbb{R} \cup \{\infty\}$ which satisfies the following three axioms:

(a) $N(c \cdot f) = \mathrm{val}(c) + N(f)$ for any $c \in K$ and $f \in K^d$,
(b) $N(f+g) \geq \min(N(f), N(g))$ for any $f, g \in K^d$,
(c) $N(f) = \infty$ if and only if $f = 0$.

We say that N is an *integral* additive norm if N takes values in $\mathbb{Z} \cup \{\infty\}$.

There is a natural bijection between lattices in K^d and integral additive norms on K^d. Namely, if N is an integral additive norm then its lattice is $\Lambda_N = \{f \in K^d : N(f) \geq 0\}$. Conversely, if Λ is any lattice in K^d then its additive norm N_Λ is given by

$$N_\Lambda(f) \;:=\; \max\{\, u \in \mathbb{Z} \,:\, z^{-u} f \in \Lambda \,\} \;=\; \min(\mathrm{val}(M^{-1} f))\,, \tag{1}$$

where M is a $d \times d$-matrix whose columns form a basis for Λ. This bijection induces a homeomorphism between the space of all additive norms (with the topology of pointwise convergence) and the space underlying the Bruhat–Tits building $\mathcal{B}_d$. In other words, non-integral additive norms can be identified with points in the simplices of $\mathcal{B}_d$.

If Λ_1 and Λ_2 are lattices then the additive norm corresponding to the intersection $\Lambda_1 \cap \Lambda_2$ is the pointwise minimum of the two norms:

$$N_{\Lambda_1 \cap \Lambda_2} \quad = \quad \min(N_{\Lambda_1}, N_{\Lambda_2})\,.$$

The pointwise maximum of two additive norms is generally not an additive norm. We write $\overline{\max}(N_{\Lambda_1}, N_{\Lambda_2})$ for the smallest norm which is pointwise greater than or equal to $\max(N_{\Lambda_1}, N_{\Lambda_2})$. Then we have

$$N_{\Lambda_1 + \Lambda_2} \quad = \quad \overline{\max}(N_{\Lambda_1}, N_{\Lambda_2})\,.$$

Our two notions of convexity on $\mathcal{B}_d$ correspond to the min and the $\overline{\max}$ of additive norms. We now present a one-dimensional example which illustrates Computational Problem A.

Example 3 (The convex hull of four 2×2-matrices)**.** Consider the following eight vectors in K^2:

$$a = \begin{pmatrix} z^{-3} \\ z^{-3} \end{pmatrix},\; b = \begin{pmatrix} z^{-4} \\ z^{5} \end{pmatrix},\; c = \begin{pmatrix} z^{3} \\ z \end{pmatrix},\; d = \begin{pmatrix} z^{-1} \\ z^{-1} \end{pmatrix},$$

$$e = \begin{pmatrix} z^{2} \\ z^{3} \end{pmatrix},\; f = \begin{pmatrix} z^{4} \\ z^{-4} \end{pmatrix},\; g = \begin{pmatrix} z \\ 1 \end{pmatrix},\; h = \begin{pmatrix} z^{4} \\ z \end{pmatrix}.$$

We compute the min-convex hull in $\mathcal{B}_2$ of the four lattices

$$\Lambda_1 = R\{a, b\},\;\; \Lambda_2 = R\{c, d\},\;\; \Lambda_3 = R\{e, f\},\;\; \Lambda_4 = R\{g, h\}.$$

The Bruhat–Tits building $\mathcal{B}_2$ is an infinite tree [1, §6.9.2], and $\mathrm{minconv}(\Lambda_1, \Lambda_2, \Lambda_3, \Lambda_4)$ is a subtree with four leaves and seven interior nodes, as shown in Figure 3. The 11 nodes in this tree represent the equivalence classes of lattices in the min-convex hull of $\Lambda_1, \Lambda_2, \Lambda_3, \Lambda_4$. Our Algorithm 2 outputs a representative lattice for each of

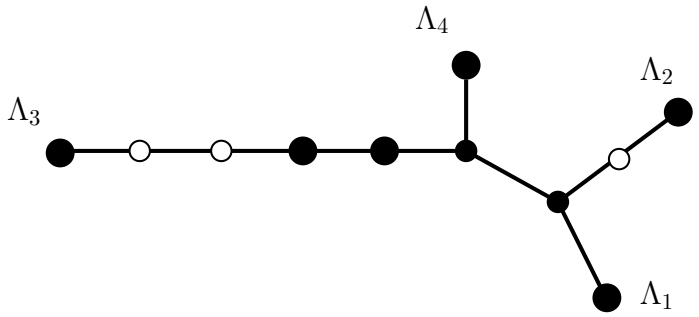


FIGURE 1. The convex hull of four points in the building $\mathcal{B}_2$.

the 11 classes:

$(1,0,7,3,6,6,5,8)$	$\{af, bf, cf, df, ef, fg, fh\}$
$(1,0,7,3,6,5,5,8)$	$\{af, bf, cf, df, ef, fg, fh\}$
$(1,0,7,3,6,4,5,8)$	$\{af, bf, cf, df, ef, fg, fh\}$
$(1,0,7,3,6,3,5,8)$	$\{af, ah, bf, bh, cf, ch, df, dh, ef, eh, fg, fh, gh\}$
$(1,0,7,3,6,2,5,7)$	$\{ac, af, ah, bc, bf, bh, cd, ce, cf, cg, ch, df, \ldots, gh\}$
$(1,0,6,3,6,1,5,6)$	$\{ac, af, ag, ah, bc, bf, bg, bh, cd, ce, cg, df, \ldots, gh\}$
$(1,0,6,3,6,1,6,6)$	$\{ag, bg, cg, dg, eg, fg, gh\}$
$(1,0,5,3,6,0,4,5)$	$\{ab, ac, ae, af, ag, ah, bc, bd, bf, bg, bh, cd, \ldots, eh\}$
$(1,1,5,3,7,0,4,5)$	$\{ab, ae, bc, bd, be, bf, bg, bh, ce, de, ef, eg, eh\}$
$(2,0,5,4,6,0,4,5)$	$\{ab, ac, ae, af, ag, ah, bd, cd, de, df, dg, dh\}$
$(3,0,5,5,6,0,4,5)$	$\{ab, ac, ae, af, ag, ah, bd, cd, de, df, dg, dh\}$

Each of the 11 lattices is represented by a vector u in $\mathbb{N}^8$ followed by a set of pairs from $\{a, b, c, d, e, f, g, h\}$. This data represents the following lattice

$$\Lambda \;=\; R\{z^{-u_1}a,\, z^{-u_2}b,\, z^{-u_3}c,\, z^{-u_4}d,\, z^{-u_5}e,\, z^{-u_6}f,\, z^{-u_7}g,\, z^{-u_8}h\},$$

in $\operatorname{minconv}(\Lambda_1, \Lambda_2, \Lambda_3, \Lambda_4)$.

Certain pairs among the eight generators form bases of $\Lambda \cong R^2$. The list of pairs indicates these bases. For example, the fourth-to-last row $(1,0,5,3,6,0,4,5)$ ... represents

$$R\{z^{-1}a, b\} = R\{z^{-1}a, z^{-5}c\} = R\{z^{-1}a, z^{-6}e\} = \cdots = R\{z^{-6}e, z^{-5}h\}.$$

The class of this lattice corresponds to the trivalent node on the right in Figure 1.

The bases can be determined from the labels of the arrows in Figure 2. A node uses a basis if and only if the node lies on the two-sided infinite path (or *apartment*) spanned by those arrows. There are eight distinct sets of pairs appearing in the above list, indicating that the tree in Figure 3 is divided into various cells. This subdivision is the key ingredient in our algorithm.

□

Returning to our general discussion, we fix an arbitrary finite subset $M = \{f_1, \ldots, f_n\}$ of K^d which spans K^d as a K-vector space, and we consider the set of

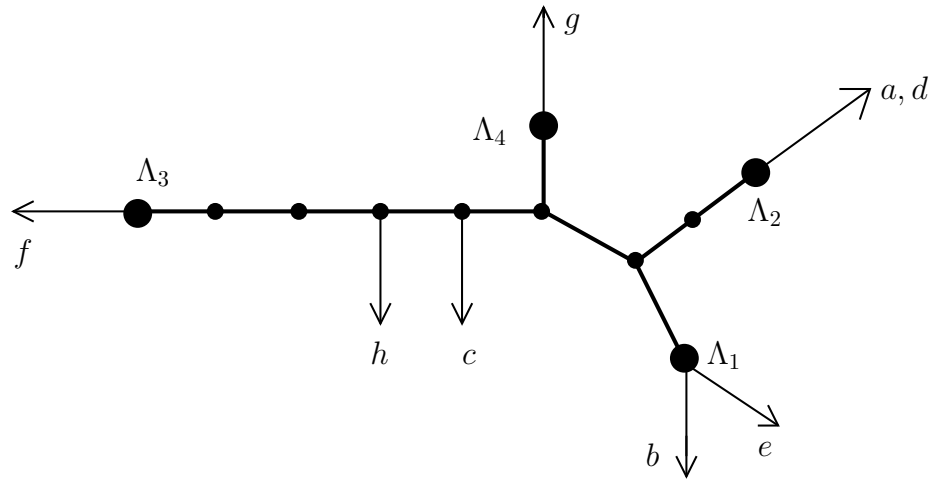


FIGURE 2. A one-dimensional membrane is an infinite tree.

all equivalence classes of lattices of the form

$$\Lambda \,=\, R\big\{z^{-u_1} f_1,\, z^{-u_2} f_2,\, z^{-u_3} f_3,\, \ldots,\, z^{-u_n} f_n\big\},$$

where $u_1, u_2, \ldots, u_n$ are any integers. This set of lattice classes is called the *membrane* spanned by M in the Bruhat–Tits building $\mathcal{B}_d$. We denote the membrane by $[M]$, and we identify it with the simplicial complex obtained by restricting $\mathcal{B}_d$ to $[M]$. If $n = d$, so that M is a basis of K^d, then the membrane $[M]$ is known as an *apartment* of the building $\mathcal{B}_d$.

Lemma 4. (Keel and Tevelev [17, Lemma 4.9]) *The membrane $[M]$ is the union of the apartments which can be formed from any d linearly independent columns of M.*

For instance, if we take $M = \{a, b, c, d, e, f, g, h\} \subset K^2$ as in Example 3, then the membrane $[M]$ is an infinite tree with seven unbounded rays, as shown in Figure 2 and derived in Example 19 below. The convex hull of $\Lambda_1 = R\{a, b\}$, $\Lambda_2 = R\{c, d\}$, $\Lambda_3 = R\{e, f\}$ and $\Lambda_4 = R\{g, h\}$ was constructed as a finite subcomplex of the infinite tree $[M]$.

The term "membrane" was coined by Keel and Tevelev [17] who showed that $[M]$ is a triangulation of the tropicalization of the subspace of K^n spanned by the rows of the $d \times n$-matrix $[f_1, \ldots, f_n]$. This result is implicit in the work of Dress and Terhalle [8, 9]. The precise statement and a self-contained proof will be given in Theorem 18 below.

The membrane $[M]$ is obviously max-convex in $\mathcal{B}_d$. However, for $d \geq 3$, membranes are generally not min-convex. Here is a simple example which shows this:

Example 5. We consider the 3×5-matrix

$$M \;=\; (f_1, f_2, f_3, f_4, f_5) \;=\; \begin{pmatrix} z & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}$$

The lattices $\Lambda_1 = R\{f_1, f_2, f_3\}$ and $\Lambda_2 = R\{f_1, f_4, f_5\}$ are in the membrane $[M]$. However, their intersection $\Lambda_1 \cap \Lambda_2 = R(0, 1, -1) + zR^3$ is a lattice which is not in $[M]$. □

While apartments and membranes are infinite subcomplexes of the Bruhat–Tits building $\mathcal{B}_d$, they have a natural finite presentation by matrices whose columns are in K^d. We can thus ask computational questions about apartments and membranes, such as:

Computational Problem B. Compute the intersection of s given apartments (or membranes) in $\mathcal{B}_d$. The input is represented by rank d matrices $M_1, \ldots, M_s$ having d rows with entries in K. The i-th apartment (or membrane) is spanned by the columns of M_i. The desired intersection is a locally finite simplicial complex of dimension $\leq d - 1$.

General solutions to Problems A and B, based on tropical convexity, will be presented in Sections 5 and 6. At this point, the reader may wish to contemplate our two problems for the special case $d = 2$: the intersection of apartments is a path which is usually finite.

Remark 6. In the theory of buildings there is another frequently used notion of convexity. Following [1, §3.6.2], it rests on the following definitions. The maximal simplices in the Bruhat–Tits building $\mathcal{B}_d$ are called *chambers*. A set $\mathcal{C}$ of chambers is *convex* if every chamber on a shortest path (in the dual graph of the simplicial complex $\mathcal{B}_d$) between two chambers of $\mathcal{C}$ also lies in $\mathcal{C}$. This notion of convexity on $\mathcal{B}_d$ agrees with convexity induced by shortest geodesics on spaces of non-positive curvature, and it is related to decompositions of semi-simple Lie groups [14]. Apartments and sub-buildings as well as intersections of convex sets are convex. A set of chambers contained in an apartment is convex if and only if it is the intersection of *roots* (or *half-apartments*). In a thick building, such as $\mathcal{B}_d$, every root is the intersection of two apartments. Hence any convex set within some apartment of $\mathcal{B}_d$ arises as the output of an algorithm for Computational Problem B. Such algorithms are our topic in Section 6. The relationship of this *classical convexity* in $\mathcal{B}_d$ to min- and max-convexity will be clarified in Proposition 20 and Theorem 29.

3. Tropical polytopes

We review the basics of tropical convexity from [7]. A subset P of $\mathbb{R}^d$ is called *tropically convex* if it is closed under linear combinations in the min-plus algebra, i.e. for any two vectors $x = (x_1, \ldots, x_d)$ and $y = (y_1, \ldots, y_d)$ in P and any scalars $\lambda, \mu \in \mathbb{R}$ we also have

$$\big(\min(x_1 + \lambda, y_1 + \mu), \ldots, \min(x_d + \lambda, y_d + \mu)\big) \;\in\; P.$$

It has become customary to write the *tropical arithmetic* operations as

$$x \oplus y \;:=\; \min(x, y) \quad \text{and} \quad x \odot y \;:=\; x + y\,.$$

In particular, if $x = (x_1, \ldots, x_d) \in P$ then $\lambda \odot x := (\lambda \odot x_1, \ldots, \lambda \odot x_d) \in P$ for all λ. Thus we can identify each tropically convex set $P \subset \mathbb{R}^d$ with its image in the *tropical projective space*, which is defined as the quotient space

$$\mathbb{TP}^{d-1} \quad := \quad \mathbb{R}^d / \mathbb{R}(1, 1, \ldots, 1)\,.$$

The canonical projection $\mathbb{R}^d \to \mathbb{TP}^{d-1}$ is denoted by ζ. The map

$$\delta(x,y) \quad := \quad \max_{1\le i<j\le d} |x_i + y_j - x_j - y_i| \tag{2}$$

from $\mathbb{R}^d \to \mathbb{R}$ is constant on products of fibers of ζ, and hence it gives rise to a map $\delta : \mathbb{TP}^{d-1} \times \mathbb{TP}^{d-1} \to \mathbb{R}$ which is a metric. We call this the *natural metric* on the tropical projective space $\mathbb{TP}^{d-1}$. The following characterizes the projection to the nearest point in a closed convex set.

Proposition 7. *Let $v \in \mathbb{R}^d$ and P be a closed tropically convex set in $\mathbb{TP}^{d-1}$. Among all vectors $w \in \mathbb{R}^d$ with $\zeta(w) \in P$ and $w \ge v$ there is a unique coordinate-wise minimal vector $\overline{w}$. The point $\zeta(\overline{w})$ minimizes the δ-distance from $\zeta(v)$ to P, and we write $\pi_P(\zeta(v)) := \zeta(\overline{w})$.*

Proof. Let $x := \zeta(v)$. Since we can always add multiples of $(1,1,\ldots,1)$ to any vector representing a point in P the set $F := \{w \in \mathbb{R}^d : \zeta(w) \in P, w \ge v\}$ is not empty. If $w, w' \in F$ then the coordinate-wise minimum $\min(w,w')$ also lies in F. Since P is closed, it then follows that the set F has a minimal element $\overline{w}$. We claim that $y := \zeta(\overline{w})$ is δ-closest to x among all points in P. Consider any point $y' \in P$. After translation we may assume $x = 0$. Of course, both y and y' are uniquely represented by non-negative vectors w and w', respectively, whose smallest coordinates are zero. So $\delta(x,y)$ is the largest coordinate of w, and $\delta(x,y')$ is the largest coordinate of w'. By construction of $\pi_P(x) := y$, we have $w_i \le w'_i$ for all i, and hence $\delta(x,y) \le \delta(x,y')$. It is clear that the value y of the map π_P for the argument x does not depend on the choice of the representative v. □

The map $\pi_P : \mathbb{TP}^{d-1} \to P$, $x \mapsto \pi_P(x)$ is the *nearest point map* onto P. Clearly, $\pi_P(x) = x$ if and only if $x \in P$. We now give an explicit formula for π_P in the special case when P is a *tropical polytope*. This means that P is the smallest tropically convex set containing a given finite collection of points $v_1, v_2, \ldots, v_n \in \mathbb{TP}^{d-1}$. Thus P is the *tropical convex hull* of these points, in symbols, $P = \mathrm{tconv}(v_1, v_2, \ldots, v_n)$.

Lemma 8. *The i-th coordinate of the nearest point map onto the tropical polytope $P = \mathrm{tconv}(v_1, v_2, \ldots, v_n)$ in $\mathbb{TP}^{d-1}$ is given by the formula*

$$\pi_P(x)_i \;=\; \min_{k\in\{1,\ldots,n\}} \max_{j\in\{1,\ldots,d\}} (v_{ki} - v_{kj} + x_j).$$

Proof. Set $y_i = \min_{k=1}^n \max_{j=1}^d (v_{ki} - v_{kj} + x_j)$. Taking $j = i$ in the maximum, we see that the vector $y = (y_1, \ldots, y_d)$ satisfies $y \ge x$. Writing $y_i = \min_{k=1}^n (\max_{j=1}^d (x_j - v_{kj}) + v_{ki})$, we find that y is a tropical linear combination of the points $v_1, \ldots, v_n$. Hence y lies in P. Moreover, y is the coordinate-wise minimal vector in $\mathbb{R}^d$ with these two properties. □

Example 9. There may be several points in a tropical polytope P which minimize the distance to a given point x. Consider the point $x = (0,1,1)$ in the plane $\mathbb{TP}^2$ and the one-dimensional polytope $P = \mathrm{tconv}((1,0,0),(0,1,0),(0,0,1))$. The projection of x onto P is $\pi_P(x) = (0,0,0) = (1,1,1)$, but

$$\delta(x,(0,0,0)) \;=\; \delta(x,(0,0,1)) \;=\; \delta(x,(0,1,0)) \;=\; 1.$$

□

The formula in Lemma 8 specifies a subdivision of the tropical polytope P into cells. These *cells* are ordinary polytopes of the special form

$$\{\, w \in \mathbb{TP}^{d-1} \,:\, w_i - w_j \leq u_{ij} \text{ for all } i \neq j \,\} \qquad (\text{for some } u_{ij} \in \mathbb{R}). \tag{3}$$

The cell containing $x \in P$ is specified by its *type*, which is the collection of index sets where "min" and "max" are attained in the identity $\pi_P(x) = x$. To be precise, we define $\text{type}(x) := (S_1, S_2, \ldots, S_d)$, where

$$\begin{aligned} S_i \;&=\; \left\{ k \in \{1,\ldots,n\} \,:\, \max_{j \in \{1,\ldots,d\}} (v_{ki} - v_{kj} + x_j) \;=\; x_i \right\} \\ &=\; \{\, k \,:\, v_{ki} - x_i \;=\; \min(v_{k1} - x_1, v_{k2} - x_2, \ldots, v_{kd} - x_d) \,\}\,. \end{aligned} \tag{4}$$

Two points of P lie in the same cell if and only if they have the same type. This subdivision of P depends on the chosen generators $v_1, v_2, \ldots, v_n$ and not just on the set P.

Remark 10. The sets $\{\, w \in \mathbb{TP}^{d-1} \,:\, w_i - w_j \,\leq\, u \,\}$ are the ordinary affine halfspaces which are also tropically convex. For integral u we call such a halfspace a *root* of $\mathbb{TP}^{d-1}$.

A point in the tropical projective space $\mathbb{TP}^{d-1}$ is a *lattice point* if it is represented by a vector x in $\mathbb{Z}^d$. We define a graph on the set of all lattice points as follows: two points x and y are connected by an edge if and only if $\delta(x, y) = 1$. The δ-distance between any two lattice points in $\mathbb{TP}^{d-1}$ is the shortest length of any path connecting these two points in the graph. A *tropical lattice polytope* is the tropical convex hull of finitely many lattice points in $\mathbb{TP}^{d-1}$. The cells of a tropical lattice polytope are intersections of roots.

Theorem 11. *The flag simplicial complex defined by this graph is a triangulation of the affine space $\mathbb{TP}^{d-1}$. It restricts to a triangulation of each cell* (3) *of each tropical lattice polytope P. We refer to this as the* standard triangulation of $\mathbb{TP}^{d-1}$, *or of P, or of* (3).

Proof. We represent points in $\mathbb{TP}^{d-1}$ by vectors with first coordinate zero. This identifies the lattice points in $\mathbb{TP}^{d-1}$ with $\mathbb{Z}^{d-1}$. The maximal simplices in the flag complex are

$$\left\{ a, a + e_{\sigma_2}, a + e_{\sigma_2} + e_{\sigma_3}, \ldots, a + e_{\sigma_2} + e_{\sigma_3} + \cdots + e_{\sigma_d} \right\},$$

where $u \in \mathbb{Z}^{d-1}$ and σ is any permutation of $\{2, \ldots, d\}$. If we fix a and let σ range over all $(d-1)!$ permutations then these simplices triangulate the unit cube with lower vertex a. Putting all these triangulated cubes together, we see that the flag complex is a triangulation of $\mathbb{TP}^{d-1}$. Each simplex in this standard triangulation is the solution set to a system of inequalities $w_i - w_j \leq u_{ij}$ where $u_{ij} + u_{ji} \leq 1$ for all $1 \leq i < j \leq d$. This implies that if w is any point in a cell (3) then that cell contains the entire simplex of the standard triangulation which has w in its relative interior. Therefore the standard triangulation of $\mathbb{TP}^{d-1}$ induces a triangulation of every tropical lattice polytope. □

Example 12 ($d = 3, n = 9$)**.** Let $v_1, v_2, \ldots, v_9$ denote the columns of

$$V \;=\; \begin{pmatrix} 0 & 0 & 0 & 1 & -3 & 1 & -3 & -4 & 0 \\ -5 & -4 & -8 & 0 & 0 & 0 & -7 & -8 & 0 \\ -3 & 2 & -3 & 0 & -2 & 2 & 0 & 0 & 0 \end{pmatrix} \tag{5}$$

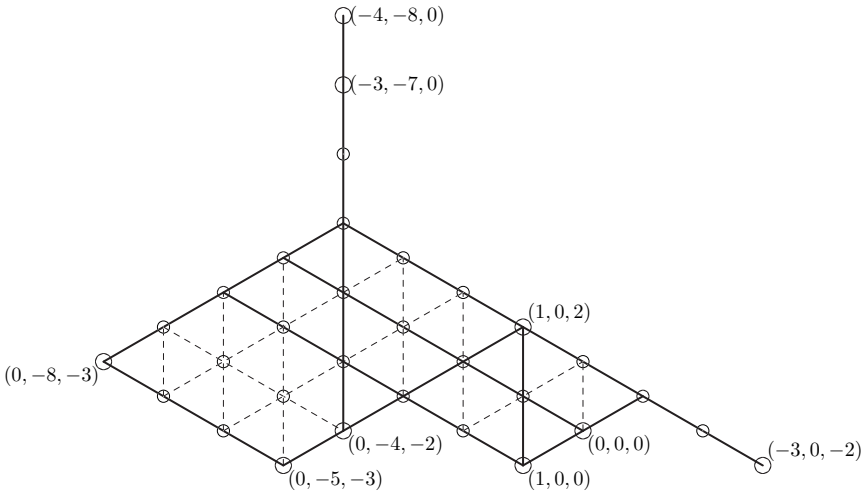


FIGURE 3. The tropical convex hull of nine labeled lattice points in $\mathbb{TP}^2$. Dashed lines and white points indicate the standard triangulation of this polygon. Solid lines and black points show the decomposition into cells (3).

We compute the tropical convex hull $P = \text{tconv}(v_1, \ldots, v_9)$ in $\mathbb{TP}^2$. The tropical lattice polygon P has ten 2-dimensional cells, 28 edges, and 19 vertices. Hence there are $10 + 28 + 19 = 57$ distinct types $\text{type}(x) = (S_1, S_2, S_3)$ among the points x in P. The standard triangulation of P is a simplicial complex with 32 triangles, 62 edges and 31 vertices, namely, the lattice points in P. It is depicted in Figure 3. □

By [7, Theorem 23], the convex hull of the rows of a matrix equals the convex hull of the columns of that same matrix. Indeed, if V is the $d \times n$-matrix whose columns are the vectors v_i then the cell complex on $P = \text{tconv}(v_1, \ldots, v_n)$ defined by the types is isomorphic to the cell complex on the convex hull in $\mathbb{TP}^{n-1}$ of the d row vectors of V.

Example 13. (Self-Duality of Tropical Polytopes) Let v_1', v_2', v_3' be the row vectors of the matrix V in (5), and let $P' = \text{tconv}(v_1', v_2', v_3')$ be their tropical convex hull in $\mathbb{TP}^8$. The tropical triangle P' contains precisely the following 31 lattice points:

$(\underline{\mathbf{4}}, \underline{\mathbf{4}}, \underline{\mathbf{4}}, \mathbf{5}, \mathbf{1}, \mathbf{5}, \mathbf{1}, \mathbf{0}, \mathbf{4})$	$(\underline{4}, \underline{4}, \underline{3}, 5, 1, 5, 1, 0, 4)$	$(\underline{4}, \underline{4}, \underline{2}, 5, 1, 5, 1, 0, 4)$
$(\underline{4}, \underline{4}, \underline{1}, 5, 1, 5, 1, 0, 4)$	$(\underline{3}, \underline{4}, \underline{3}, 5, 1, 5, 1, 0, 4)$	$(\underline{3}, \underline{4}, \underline{2}, 5, 1, 5, 1, 0, 4)$
$(\underline{3}, \underline{4}, \underline{1}, 5, 1, 5, 1, 0, 4)$	$(\underline{3}, \underline{4}, \underline{0}, 5, 1, 5, 1, 0, 4)$	$(\underline{2}, \underline{4}, \underline{2}, 5, 1, 5, 1, 0, 4)$
$(\underline{2}, \underline{4}, \underline{1}, 5, 1, 5, 1, 0, 4)$	$(\underline{2}, \underline{4}, \underline{0}, 5, 1, 5, 1, 0, 4)$	$(\underline{3}, \underline{4}, 0, \underline{6}, 2, 6, 1, 0, 5)$
$(\underline{3}, 4, 0, \underline{7}, \underline{3}, 7, 1, 0, 6)$	$(\underline{3}, 4, 0, \underline{8}, \underline{4}, 8, 1, 0, 7)$	$(\underline{1}, \underline{4}, \underline{1}, 4, 1, 5, 1, 0, 4)$
$(\underline{1}, \underline{4}, \underline{0}, 4, 1, 5, 1, 0, 4)$	$(\underline{2}, \underline{4}, 0, 5, \underline{2}, 6, 1, 0, 5)$	$(\underline{3}, \underline{4}, 0, 6, \underline{3}, 7, 1, 0, 6)$
$(\underline{3}, 4, 0, \underline{7}, \underline{4}, 8, 1, 0, 7)$	$(\underline{3}, 4, 0, \underline{8}, \underline{5}, 8, 1, 0, 8)$	$(\underline{0}, \underline{4}, \underline{0}, 3, 1, 5, 1, 0, 3)$
$(\underline{1}, \underline{4}, 0, 4, \underline{2}, 6, 1, 0, 4)$	$(\underline{2}, \underline{4}, 0, 5, \underline{3}, 7, 1, 0, 5)$	$(\underline{3}, \underline{4}, 0, 6, \underline{4}, 8, 1, 0, 6)$
$(\underline{3}, 4, 0, \underline{7}, \underline{5}, 8, 1, 0, 7)$	$(\underline{3}, 4, 0, \underline{8}, \underline{6}, 8, 1, 0, 8)$	$(\underline{3}, 4, 0, \underline{8}, \underline{7}, 8, 1, 0, 8)$
$(\underline{\mathbf{3}}, \mathbf{4}, \mathbf{0}, \underline{\mathbf{8}}, \underline{\mathbf{8}}, \mathbf{8}, \mathbf{1}, \mathbf{0}, \mathbf{8})$	$(\underline{0}, \underline{5}, \underline{0}, 3, 1, 5, 2, 1, 3)$	$(\underline{0}, 5, \underline{0}, 3, 1, 5, \underline{3}, 2, 3)$
$(\underline{\mathbf{0}}, \mathbf{5}, \underline{\mathbf{0}}, \mathbf{3}, \mathbf{1}, \mathbf{5}, \mathbf{3}, \underline{\mathbf{3}}, \mathbf{3})$		

Here each point is represented uniquely by a non-negative vector with a zero entry. The boldfaced vectors represent the given points $v_1', v_2', v_3' \in \mathbb{TP}^8$. The underlined triples of coordinates will be explained in Example 23. The tropical triangle P', which lives in $\mathbb{TP}^8$, is isomorphic to the tropical 9-gon P of Example 12, which lives in $\mathbb{TP}^2$ and is depicted in Figure 3. According to equation (14) in [7, page 16], the isomorphism between the two tropical polygons is given by the piecewise-linear maps

$$\begin{aligned} P \to P',\ (x_1,x_2,x_3) &\mapsto \big(\overset{3}{\underset{i=1}{\min}}(v_{i1}-x_i),\ldots,\overset{3}{\underset{i=1}{\min}}(v_{i9}-x_i)\big)\,, \\ P' \to P,\ (y_1,\ldots,y_9) &\mapsto \big(\min_j(v_{1j}-y_j),\ldots,\min_j(v_{3j}-y_j)\big)\,. \end{aligned} \tag{6}$$

These bijections are inverses of each other. They are linear on each cell, and they identify the types: if $x \in P$ and $\text{type}(x) = (S_1, S_2, S_3)$ then the corresponding point $y \in P'$ has $\text{type}(y) = (S_1', S_2', \ldots, S_9')$ where $S_j' = \{i : j \in S_i\}$. The 31 lattice points in $\mathbb{TP}^8$ that are listed above get sent to the 31 lattice points in Figure 3 by the map $P' \to P$. □

We close with the remark that several algorithms are available for computing a tropical polytope P from its defining matrix $V = (v_{ij})$. They will be discussed in Section 5.

4. Tropical linear spaces and membranes

This section is concerned with the relationship between tropical linear spaces, valuated matroids [8, 9], and membranes [17] in the Bruhat–Tits building. In order to think of these objects as tropical polytopes, we shall now augment the real numbers $\mathbb{R}$ by the extra element ∞. Note that ∞ is the additively neutral element in the min-plus algebra. We define the *compactified tropical projective space* $\overline{\mathbb{TP}}^{d-1}$ to be $(\mathbb{R}\cup\{\infty\})^d \setminus \{(\infty,\ldots,\infty)\}$ modulo the equivalence relation given by tropical scalar multiplication. The notions of tropical convexity, tropical polytopes and lattice points make sense in $\overline{\mathbb{TP}}^{d-1}$. When extending the metric δ to $\overline{\mathbb{TP}}^{d-1}$ we use the convention that $\infty - \infty = 0$ in the formula (2). Proposition 7 and Lemma 8 remain valid, and there is a standard triangulation of $\overline{\mathbb{TP}}^{d-1}$. That standard triangulation coincides with the compactified apartment in the work of Werner [25, 26]. We also refer to Alessandrini [2] whose tropical approach to buildings is similar to ours and is aimed at applications in Teichmüller theory.

For experts on buildings we note that our two notions of convexity in Problem A reflect two different compactifications of the Bruhat–Tits buildings $\mathcal{B}_d$. The first is featured in [18, 25] and we call it the *max-compactification*. It is a simplicial complex whose vertices are all free R-submodules of K^d, and the boundary consists of modules of rank less than d. The second compactification, which we call the *min-compactification*, arises more naturally from tropical geometry. Its points consist of all additive seminorms on K^d. An additive seminorm is a function $N : K^d \to \mathbb{R} \cup \{\infty\}$ which satisfies the first two axioms of an additive norm. If N is an additive seminorm then $N^{-1}(\infty)$ is a linear subspace of K^d. The boundary of the min-compactification consists of additive seminorms for which $N^{-1}(\infty)$ is positive-dimensional. We shall not dwell on the matters here, but we do wish to underline that our combinatorial results are compatible with these compactifications.

We now review the definition of tropical linear spaces [22, 23]. Fix two positive integers $d \leq n$ and consider a map $p : \{1, 2, \ldots, n\}^d \to \mathbb{R} \cup \{\infty\}$. Following Dress and Terhalle [8, 9], we say that p is a *valuated matroid* if $p(\omega)$ depends only on the unordered set $\{\omega_1, \ldots, \omega_d\}$, and $p(\omega) = \infty$ whenever ω has fewer than d elements, and p satisfies the following variant of the basis exchange axiom: for any $(d-1)$-subset σ and any $(d+1)$-subset τ of $\{1, 2, \ldots, n\}$, the minimum of the list of numbers $\big(p(\sigma \cup \tau_i) + p(\tau \setminus \{\tau_i\}) \,:\, i = 1, 2, \ldots, d+1\big)$ is attained at least twice. This axiom is equivalent to saying that p lies in the *tropical prevariety* [20] specified by the set of all *quadratic Plücker relations*.

Fix a valuated matroid p. The associated *tropical linear space* L_p consists of all points $x \in \overline{\mathbb{TP}}^{n-1}$ such that, for any $(d+1)$-subset τ of $\{1, 2, \ldots, n\}$, the minimum of the numbers $p(\tau \backslash \{\tau_i\}) + x_{\tau_i}$, for $i = 1, 2, \ldots, d$, is attained at least twice. This list of numbers represents a *circuit* of p. The tropical linear space L_p is tropically convex, and it can be represented as a tropical lattice polytope as follows. For any $(d-1)$-subset σ of $\{1, \ldots, n\}$ let $p(\sigma *)$ denote the vector in $(\mathbb{R} \cup \{\infty\})^n$ whose j-th coordinate equals $p(\sigma \cup \{j\})$. We regard $p(\sigma *)$ as a point in $\overline{\mathbb{TP}}^{n-1}$, or, combinatorially, as a *cocircuit* of the valuated matroid p.

Theorem 14. (Yu and Yuster [27, Theorem 16]) *The tropical linear space L_p is the tropical convex hull in the compactified tropical projective space $\overline{\mathbb{TP}}^{n-1}$ of all the cocircuits $p(\sigma *)$ of the underlying valuated matroid $p : \{1, 2, \ldots, n\}^d \to \mathbb{R} \cup \{\infty\}$.*

The tropical linear space L_p is tropically convex. Hence it has a nearest point map π_{L_p} which takes any point $x \in \overline{\mathbb{TP}}^{n-1}$ to the coordinate-wise minimum in $\{w \in L_p : w \geq x\}$. We now present two rules for evaluating this map.

The Blue Rule. Form the vector $w \in \mathbb{R}^n$ whose coordinates are

$$w_i \quad = \quad \min_{\sigma} \max_{j \notin \sigma} \big(p(\sigma \cup \{i\}) - p(\sigma \cup \{j\}) + x_j \big). \tag{7}$$

Here the minimum is over all $(d-1)$-subsets σ of $\{1, 2, \ldots, n\}$.

The Red Rule. Start with $v = (0, 0, \ldots, 0)$. For each $(d+1)$-set τ do: If the minimum of the numbers $p(\tau \backslash \{\tau_i\}) + x_{\tau_i}$ is attained only once, for the index i, then let $\gamma_{\tau,i}$ be the difference of the second smallest number minus that minimum, set $v_{\tau_i} := \max(v_{\tau_i}, \gamma_{\tau,i})$, and iterate.

The terms *Blue Rule* and *Red Rule* were introduced by Ardila [3]. The following theorem extends his main result in [3] from ordinary matroids to valuated matroids:

Theorem 15. *Let p be a valuated matroid, L_p its tropical linear space and $x \in \overline{\mathbb{TP}}^{n-1}$. If v and w are computed by the Red Rule and the Blue Rule then $\pi_{L_p}(x) = x + v = w$.*

Sketch of Proof. In the case of ordinary matroids, the image of p lies in $\{0, \infty\}$. This special case is [3, Theorem 1]. Ardila's proof easily generalizes to valuated matroids. The correctness of the Blue Rule also follows from Lemma 8 and Theorem 14. □

Remark 16. The Red Rule and the Blue Rule produce the identical result in the special case when $x = (0, 0, \ldots, 0)$. We find that $\pi_P(0, 0, \ldots, 0) \in L_p$ is the tropical sum of all cocircuits $p(\sigma *)$ of the valuated matroid p, provided each cocircuit is represented by the unique vector whose coordinates are non-negative and has at least one coordinate zero.

We now apply tropical convexity to the Bruhat–Tits building $\mathcal{B}_d$. We begin with a review on how tropical linear spaces are related to ordinary linear spaces over the field $K = \mathbb{C}(\!(x)\!)$. Let M be a $d \times n$-matrix of rank d with entries in K. The row space of M is a d-dimensional linear subspace of K^n, or a $(d-1)$-dimensional subspace of the projective space $\mathbb{P}_K^{n-1}$. If ω is an ordered list of d elements in $\{1, 2, \ldots, n\}$ then M_ω denotes the corresponding $d \times d$-submatrix. The matrix M defines a valuated matroid p by the rule

$$p(\omega) \quad = \quad \mathrm{val}\big(\det(M_\omega)\big)\,. \tag{8}$$

Note that $p(\omega) = \infty$ if and only if M_ω is not invertible over K.

Proposition 17. (Speyer and Sturmfels [23, Theorem 2.1]) *The lattice points in the tropical linear space L_p are precisely the points* $\mathrm{val}(v)$ *where v is in the row space of M.*

Since L_p is a tropical lattice polytope, the standard triangulation of $\overline{\mathbb{TP}}^{n-1}$ restricts to a triangulation of L_p. We shall present a self-contained proof of the following result.

Theorem 18. (Keel and Tevelev [17, Theorem 4.11]) *Let $M = (f_1, f_2, \ldots, f_n)$ be a $d \times n$-matrix of rank d over K, and let L_p be the associated tropical linear space. Then*

$$\Psi_M \,:\, R\big\{z^{-u_1} f_1,\, z^{-u_2} f_2,\, \ldots,\, z^{-u_n} f_n\big\} \,\mapsto\, \pi_{L_p}(u_1, u_2, \ldots, u_n)$$

is a well-defined map, and it induces an isomorphism of simplicial complexes between the membrane $[M]$ and the standard triangulation of L_p.

Proof. Consider any lattice $\Lambda = R\big\{z^{-u_1} f_1,\, z^{-u_2} f_2,\, \ldots,\, z^{-u_n} f_n\big\}$ in the membrane, and set $(v_1, v_2, \ldots, v_n) = \pi_{L_p}(u_1, u_2, \ldots, u_n)$. We claim that

$$v_i \quad = \quad \max\{\, \mu \in \mathbb{Z} \,:\, z^{-\mu} f_i \in \Lambda \,\}. \tag{9}$$

We first prove the inequality "$\leq$". By the Red Rule in Theorem 15, we have $v_i = \gamma_{\tau,i} + u_i$ for some $(d+1)$-set τ containing i. We may assume $\tau_{d+1} = i$. Then $\{f_{\tau_1}, \ldots, f_{\tau_d}\}$ is a basis of K^d, and we can write

$$f_i \quad = \quad p_1 f_{\tau_1} + p_2 f_{\tau_2} + \cdots + p_d f_{\tau_d} \qquad \text{for some } p_1, \ldots, p_d \in K.$$

Our choice of the $(d+1)$-set τ in the Red Rule means that

$$u_i + \gamma_{\tau,i} \quad = \quad \min\{\, \mathrm{val}(p_j) + u_{\tau_j} \,:\, j = 1, 2, \ldots, d \,\} \quad \geq \quad 0\,,$$

and therefore

$$f_i z^{-u_i - \gamma_{\tau,i}} \;=\; p_1 z^{u_{\tau_1}} (f_{\tau_1} z^{-u_{\tau_1}}) + \cdots + p_d z^{u_{\tau_d}} (f_{\tau_d} z^{-u_{\tau_d}}) \;\in\; \Lambda\,. \tag{10}$$

This proves the inequality "$\leq$". The converse "$\geq$" holds because $z^{-\mu} f_i$ lies in Λ if and only it lies in the R-submodule spanned by d of the n generators, and a representation (10) is the only way this can happen. Indeed, by Lemma 4, the membrane $[M]$ is the union of the apartments $[(f_{\tau_1}, \ldots, f_{\tau_d})]$ that can be formed from the d-subsets $\tau \subseteq \{1, 2, \ldots, n\}$.

The identity (9) shows that the map Ψ_M which takes the lattice $R\big\{z^{-u_1} f_1,$ $\ldots,\, z^{-u_n} f_n\big\}$ to the point $\pi_{L_p}(u_1, \ldots, u_n)$ is well-defined, and is a bijection between the membrane $[M]$ and the lattice points in the tropical linear space L_p. This bijection takes adjacent lattices to points of δ-distance one in L_p and conversely. Hence it induces an isomorphism between the flag simplicial complexes of these two graphs. $\square$

Example 19. Let $d = 2$, $n = 8$ and let $M = (a, b, c, d, e, f, g, h)$ be as in Example 3. The valuated matroid p of the matrix M maps pairs of columns to $\mathbb{Z} \cup \{\infty\}$ as follows:

$$\begin{pmatrix} aa & ab & ac & \cdots & ah \\ ab & bb & bc & \cdots & bh \\ ac & bc & cc & \cdots & ch \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ ah & bh & ch & \cdots & hh \end{pmatrix} \mapsto \begin{pmatrix} \infty & -7 & -2 & \infty & -1 & -7 & -3 & -2 \\ -7 & \infty & -3 & -5 & -1 & -8 & -4 & -3 \\ -2 & -3 & \infty & 0 & 3 & -1 & 2 & 4 \\ \infty & -5 & 0 & \infty & 1 & -5 & -1 & 0 \\ -1 & -1 & 3 & 1 & \infty & -2 & 2 & 3 \\ -7 & -8 & -1 & -5 & -2 & \infty & -3 & 0 \\ -3 & -4 & 2 & -1 & 2 & -3 & \infty & 2 \\ -2 & -3 & 4 & 0 & 3 & 0 & 2 & \infty \end{pmatrix}$$

The rows of the 8×8-matrix to the right are the cocircuits $p(\sigma*)$ of the valuated matroid p. They represent seven distinct points in $\overline{\mathbb{TP}}^7$ (rows 1 and 4 give the same point). The tropical linear space L_p is the tropical convex hull of these seven points in $\overline{\mathbb{TP}}^7$. This convex hull is the tree depicted in Figure 2. A systematic algorithm for drawing such a tree L_p, given its valuated matroid p, is the *neighbor-joining method* from phylogenetics; see [23, §6]. □

Theorem 18 states that every lattice point $(u_1, \ldots, u_n)$ in L_p uniquely represents a lattice $\Lambda_u = R\{z^{-u_1} f_1, \ldots, z^{-u_n} f_n\}$ in the membrane $[M]$. The lattice Λ_u specifies a matroid M_u of rank d on $\{1, 2, \ldots, n\}$. This is an ordinary (not valuated) matroid. The bases of M_u are the sets $\{\tau_1, \ldots, \tau_d\}$ such that $\{z^{-u_{\tau_1}} f_{\tau_1}, \ldots, z^{-u_{\tau_d}} f_{\tau_d}\}$ spans the lattice Λ. The matroid M_u can be read off directly from the valuated matroid p as follows: its bases are the d-sets τ such that the expression $p(\tau) - u_{\tau_1} - \cdots - u_{\tau_d}$ is minimal. The set of all matroids M_u, as u ranges over the tropical linear space L_p, forms a *matroid subdivision* of the matroid polytope of the matrix M over the field K. This is the identification of tropical linear spaces with matroid subdivisions as studied in [16, 22].

Our algorithm for Computational Problem A in Section 5 will output each lattice Λ_u in the min-convex hull as a pair (u, M_u), where u is a point in a tropical linear space L_p and M_u is a matroid. We saw this format already in Example 3. For instance, consider the point $u = (2, 0, 5, 4, 6, 0, 4, 5)$ listed there. It lies in the tropical line L_p of Example 19. The rank 2 matroid M_u has the set of bases $\{ab, ac, ae, af, ag, ah, bd, cd, de, df, dg, dh\}$.

The classical notion of convexity in buildings in Remark 6 is related to tropical convexity as follows. For a chamber C in $\mathcal{B}_d$ let $\mathrm{vert}(C)$ be its set of vertices. Now consider a set $\mathcal{C}$ of chambers contained in some apartment $\mathcal{A}$. We identify $\mathcal{A}$ with $\mathbb{TP}^{d-1}$ and we note that the classical notion of a *root* (or *half-apartment*) of $\mathcal{A}$ agrees with our definition of a root in $\mathbb{TP}^{d-1}$ from Remark 10. We consider the following set of lattice points in $\mathbb{TP}^{d-1}$:

$$\mathrm{vert}(\mathcal{C}) \;:=\; \bigcup \{\, \mathrm{vert}(C) \,:\, C \in \mathcal{C} \,\}$$

Our next result holds because the convex subsets of chambers in $\mathcal{A}$ are intersections of roots, or equivalently, intersections of $\mathcal{A}$ with other apartments. See also Theorem 29.

Proposition 20. *A finite set $\mathcal{C}$ of chambers in an apartment $\mathcal{A} \cong \mathbb{TP}^{d-1}$ is convex if and only if* $\mathrm{vert}(\mathcal{C})$ *is the set of lattice points in a tropical lattice polytope of the form (3).*

Proposition 20 implies that the convex sets of chambers are precisely the maximal simplices in the standard triangulation of those tropical lattice polytopes which are at the same time (possibly unbounded) ordinary convex polyhedra. In other words, Proposition 20 holds verbatim for infinite $\mathcal{C}$ if $\mathbb{TP}^{d-1}$ is replaced by its compactification $\overline{\mathbb{TP}}^{d-1}$.

5. Convex hulls in the Bruhat–Tits building

In this section and the next we present algorithmic implications of the theory developed so far. We begin with Computational Problem A: how to find min-convex hulls in $\mathcal{B}_d$. The input is a list of s invertible $d \times d$-matrices $M_1, M_2, \ldots, M_s$ with entries in the field $K = \mathbb{C}(\!(z)\!)$, each representing the equivalence class of its column lattice $\Lambda_i = \text{image}_R(M_i)$.

5.1. **The retraction of min-convex hulls to a membrane.** Let $M = (f_1, \ldots, f_n)$ be any matrix in $K^{d\times n}$ of rank d and let $[M]$ be the membrane in $\mathcal{B}_d$ which is spanned by the n column vectors of M. There is a natural retraction r_M from $\mathcal{B}_d$ onto $[M]$ given by

$$r_M : \Lambda \mapsto (\Lambda \cap K\{f_1\}) + \cdots + (\Lambda \cap K\{f_n\}) \tag{11}$$

This map restricts to the identity on the membrane $[M]$.

Let V be the d-dimensional subspace of K^n spanned by the rows of M, and let p be its valuated matroid as in formula (8). By Proposition 17, the tropicalization of the classical linear space V over the field K equals the tropical linear space L_p. The map Ψ_M in Theorem 18 allows us to identify the lattice points in L_p with the membrane $[M]$.

Lemma 21. *Fix a membrane $[M]$ in $\mathcal{B}_d$ and consider any lattice $\Lambda = \text{image}_R(M_0)$ where $M_0 \in GL_K(d)$. Then the following three lattice points in $\overline{\mathbb{TP}}^{n-1}$ coincide:*

(a) *$\Psi_M(r_M(\Lambda))$, where Ψ_M is the bijection of Theorem 18 between $[M]$ and the lattice points in L_p,*
(b) *$(N_\Lambda(f_1), \ldots, N_\Lambda(f_n))$, where N_Λ is the integral additive norm corresponding to Λ,*
(c) *the tropical sum (coordinatewise minimum) of the rows of the matrix $\text{val}(M_0^{-1} \cdot M)$.*

Proof. The equivalence of (a) and (b) follow from the definitions of N_Λ and r_M, and from equation (9). The equivalence of (b) and (c) follows from equation (1). □

As a consequence, we get the following explicit description of the retraction of a min-convex hull onto a membrane. This establishes the correctness of Algorithm 1 below.

Proposition 22. *Let $\Lambda_1, \Lambda_2, \ldots, \Lambda_s$ be the lattices spanned by the columns of the matrices $M_1, M_2, \ldots, M_s \in GL_d(K)$. Let $[M]$ be any membrane in $\mathcal{B}_d$. The simplicial complex*

$$r_M(\text{minconv}(\Lambda_1, \Lambda_2, \ldots, \Lambda_s)) \subset [M]$$

coincides with the standard triangulation of the tropical polytope

$$\text{tconv}\big(\Psi_M(r_M(\Lambda_1)), \Psi_M(r_M(\Lambda_2)), \ldots, \Psi_M(r_M(\Lambda_s))\big) \;\subset\; L_p \;=\; \text{val}(\text{kernel}(M)).$$

Proof. By the definition of the integral additive norm N_Λ in formula (1), we have

$$N_{(z^{-a}\Lambda)\cap(z^{-a'}\Lambda')} \;=\; \min\left(a + N_\Lambda, a' + N_{\Lambda'}\right).$$

By Lemma 21, for any integers $a_1, a_2, \ldots, a_s$, the image under the map Ψ_M of the retraction $r_M(z^{-a_1}\Lambda_1 \cap \cdots \cap z^{-a_s}\Lambda_s)$ coincides with the tropical linear combination

$$(a_1 \odot \Psi_M(r_M(\Lambda_1))) \oplus \cdots \oplus (a_s \odot \Psi_M(r_M(\Lambda_s))).$$

The simplicial complex structure of $[M]$ coincides with the standard triangulation of the tropical linear space L_p, which induces the simplicial complex structure on the lattice points in the tropical polytope. Hence the retraction of the min-convex hull onto the membrane coincides with the standard triangulation of the tropical polytope. □

Input: matrices $M_1, \ldots, M_s \in \mathrm{GL}_d(K)$ and a $d \times n$ matrix M over K with rank d
Output: retraction $r_M(\mathrm{minconv}(\Lambda_1, \ldots, \Lambda_s))$ onto the membrane $[M]$, where $\Lambda_i = \mathrm{image}_R(M_i)$ for $i = 1, \ldots, s$.
for $i \leftarrow 1, 2, \ldots, s$ **do**
  $\Psi_M(r_M(\Lambda_i)) \leftarrow$ tropical sum of the rows of $\mathrm{val}(M_i^{-1} \cdot M)$
return $\mathrm{tconv}(\Psi_M(r_M(\Lambda_1)), \Psi_M(r_M(\Lambda_2)), \ldots, \Psi_M(r_M(\Lambda_s)))$

Algorithm 1: Retraction of a min-convex hull in $\mathcal{B}_d$ onto a given membrane.

Example 23. (Illustration of Algorithm 1) We consider the three lattices $\Lambda_1, \Lambda_2, \Lambda_3$ in the Bruhat–Tits building $\mathcal{B}_3$ which are represented by the invertible 3×3-matrices

$$M_1 = \begin{pmatrix} 1 & z^5 & z^{-3} \\ z^4 & z & z^{-3} \\ z^{-3} & z^2 & z^{-3} \end{pmatrix},\; M_2 = \begin{pmatrix} z^2 & z^{-2} & z^2 \\ z^3 & z^5 & z^5 \\ 1 & 1 & z^4 \end{pmatrix},\; M_3 = \begin{pmatrix} z^2 & z^{-1} & z \\ z^{-2} & z^{-3} & z^3 \\ z^3 & z & 1 \end{pmatrix}.$$

Set $M := (M_1, M_2, M_3)$. Then the vectors $\Psi_M(r_M(\Lambda_1))$, $\Psi_M(r_M(\Lambda_2))$, and $\Psi_M(r_M(\Lambda_3))$ are the precisely the rows of the 3×9-matrix V in (5). That matrix was analyzed in Examples 12 and 13. Hence the tropical convex hull (of the rows) of V is the tropical polygon P in Figure 3.

The 31 lattices in P are encoded by the 31 lattice points in Figure 3, or by the 31 lattice vectors listed in Example 13. If $u = (u_1, u_2, \ldots, u_9) \in \mathbb{Z}^9$ is one these vectors then the corresponding lattice $\Lambda \subset K^3$ is generated by the nine columns of the 3×9-matrix

$$M \cdot \mathrm{diag}(z^{-u_1}, z^{-u_2}, \ldots, z^{-u_9}).$$

The underlined coordinates of u give the lexicographically first basis $\{i, j, k\}$ of the matroid M_u. This writes Λ as the column lattice of the matrix $M \cdot \mathrm{diag}(z^{-u_i}, z^{-u_j}, z^{-u_k})$. □

5.2. **Computing min-convex hulls in $\mathcal{B}_d$.** Algorithm 1 would compute the min-convex hull in $\mathcal{B}_d$ if we input a membrane that contains it. Algorithm 2 below iteratively finds such a membrane, starting from the membrane $[M]$ spanned by the given generators of $\Lambda_1, \ldots, \Lambda_s$. The idea is to compute the retraction P of the min-convex hull onto $[M]$, to identify the fiber over every lattice in P, and then to enlarge our membrane by the fibers.

As seen in the proof of Proposition 22, each lattice in the desired convex hull,

$$z^{-a_1}\Lambda_1 \cap \cdots \cap z^{-a_s}\Lambda_s \ \in \ \mathrm{minconv}(\Lambda_1, \ldots, \Lambda_s),$$

is mapped by the composition $\Psi_M \circ r_M$ to the tropical linear combination

$$a_1 \odot \Psi_M(r_M(\Lambda_1)) \oplus \cdots \oplus a_s \odot \Psi_M(r_M(\Lambda_s)) \ \in \ P.$$

Our aim is to list all lattices in the fiber $\{\Lambda \in \mathrm{minconv}(\Lambda_1, \ldots, \Lambda_s) : \Psi_M(r_M(\Lambda)) = v\}$ over a lattice point $v \in P$. There are infinitely many ways to write v as an integer tropical linear combination of $\Psi_M(r_M(\Lambda_1)), \ldots, \Psi_M(r_M(\Lambda_s))$. However, since the min-convex hulls in B_d are finite, the fibers under the retraction are finite, too. We can make sure that the loop in step 2 is finite, as follows. For a fixed $v \in P$, let C_v be the set of coefficients $a \in \mathbb{Z}^s$ such that $v = \bigoplus_{i=1}^s (a_i \odot \Psi_M(r_M(\Lambda_i)))$. Then C_v is a partially ordered set with $a \leq b$ in C_v if $a_i \leq b_i$ for all $i = 1, \ldots, s$. This partial order is compatible with the inclusion order on the fiber, i.e. $a \leq b$ implies $\bigcap_{i=1}^s (z^{-a_i}\Lambda_i) \subseteq \bigcap_{i=1}^s \left(z^{-b_i}\Lambda_i\right)$. Note that if $a, b \in C_v$ then $a \oplus b \in C_v$, so there is a unique minimal element in C_v. Starting from the unique minimal element in C_v, we do a finite depth-first-search on the Hasse diagram of C_v to enumerate the fiber over v. At every step, we increment a coordinate by 1 if the new lattice is strictly larger. Otherwise, further incrementing that coordinate will not give us new lattices in the fiber, so we abandon that branch and backtrack. In this manner we reach all elements in the fiber without going through an infinite loop. As a byproduct, Algorithm 2 produces a membrane $[M']$ which contains the min-convex hull.

Input: matrices $M_1, M_2, \ldots, M_s \in \mathrm{GL}_d(K)$
Output: $\mathrm{minconv}(\Lambda_1, \ldots, \Lambda_s)$ in $\mathcal{B}_d$, where $\Lambda_i = \mathrm{image}_R(M_i)$
$M \leftarrow (M_1, \ldots, M_s) \in K^{d \times ds}$
$M' \leftarrow M$
$P \leftarrow r_M(\mathrm{minconv}(\Lambda_1, \ldots, \Lambda_s))$, computed by Algorithm 1
1 **foreach** lattice point $v \in P$ **do**
  $\Lambda \leftarrow R\{z^{-v_j} f_j\}$ where f_j is the j^{th} column of M
2 **foreach** $a \in \mathbb{Z}^s$ such that $v = \bigoplus_{i=1}^s (a_i \odot \Psi_M(r_M(\Lambda_i)))$ **do**
    if $\Lambda \subsetneq \bigcap_{i=1}^s (z^{-a_i}\Lambda_i)$ **then**
      Augment the columns of M' with minimal generators of $\bigcap_{i=1}^s (z^{-a_i}\Lambda_i)$ that are not in Λ.
$P' \leftarrow r_{M'}(\mathrm{minconv}(\Lambda_1, \ldots, \Lambda_s))$, computed by Algorithm 1
return P'

Algorithm 2: Min-convex hull in the Bruhat–Tits building $\mathcal{B}_d$.

Example 24. We illustrate Algorithm 2 by computing the min-convex hull of three points in the Bruhat–Tits building $\mathcal{B}_3$. The input points are given by the three invertible matrices

$$M_1 = \begin{pmatrix} z & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \ M_2 = \begin{pmatrix} z & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \ M_3 = \begin{pmatrix} z & 1 & 4 \\ 0 & 2 & 5 \\ 0 & 3 & 6 \end{pmatrix}.$$

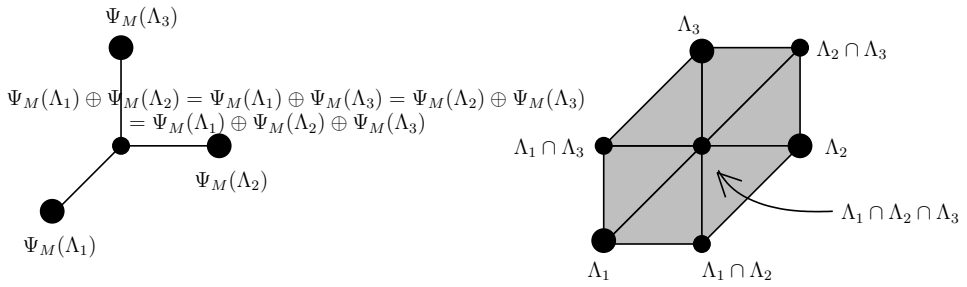


FIGURE 4. The two iterations of Algorithm 2 for $\Lambda_1, \Lambda_2, \Lambda_3$ as in Example 24.

We start with the membrane spanned by $M = (M_1, M_2, M_3)$, and hence with

$$\begin{pmatrix} \Psi_M(r_M(\Lambda_1)) \\ \Psi_M(r_M(\Lambda_2)) \\ \Psi_M(r_M(\Lambda_3)) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & -1 & -1 & 0 & -1 & -1 \\ 0 & -1 & -1 & 0 & 0 & 0 & 0 & -1 & -1 \\ 0 & -1 & -1 & 0 & -1 & -1 & 0 & 0 & 0 \end{pmatrix} \tag{12}$$

The tropical convex hull of these three row vectors has precisely one more lattice point:

$$\begin{aligned} v &= (0,-1,-1,0,-1,-1,0,-1,-1) \\ &= \Psi_M(r_M(\Lambda_1)) \oplus \Psi_M(r_M(\Lambda_2)) \\ &= \Psi_M(r_M(\Lambda_1)) \oplus \Psi_M(r_M(\Lambda_3)) \\ &= \Psi_M(r_M(\Lambda_2)) \oplus \Psi_M(r_M(\Lambda_3)) \\ &= \Psi_M(r_M(\Lambda_1)) \oplus \Psi_M(r_M(\Lambda_2)) \oplus \Psi_M(r_M(\Lambda_3)). \end{aligned}$$

The set C_v consists of the vectors $(0,0,a), (0,b,0)$ and $(c,0,0)$ where $a, b, c \in \mathbb{N}$. The unique minimal element is $(0,0,0)$. As its corresponding lattice $zR^3 = \Lambda_1 \cap \Lambda_2 \cap \Lambda_3$ lies in $[M]$, this point adds no new columns to M'. Since $\Lambda_1 \cap \Lambda_2 \cap z^{-1}\Lambda_3 = \Lambda_1 \cap \Lambda_2 \cap z^{-2}\Lambda_3$, all lattices $\Lambda_1 \cap \Lambda_2 \cap z^{-a}\Lambda_3$ are identical for $a \geq 1$. So we can abandon the branch $(0,0,a)$ in C_v after $(0,0,1)$. Similarly, we only need to consider up to $(0,1,0)$ and $(1,0,0)$.

After comparing zR^3 with the lattices $\Lambda_1 \cap \Lambda_2 \cap z^{-1}\Lambda_3$, $\Lambda_1 \cap z^{-1}\Lambda_2 \cap \Lambda_3$ and $z^{-1}\Lambda_1 \cap \Lambda_2 \cap \Lambda_3$ respectively, we augment the columns of M' with the three vectors:

$$\begin{aligned} (0,1,-1) &\in (\Lambda_1 \cap \Lambda_2 \cap z^{-1}\Lambda_3) \setminus zR^3 \\ (0,1,2) &\in (\Lambda_1 \cap z^{-1}\Lambda_2 \cap \Lambda_3) \setminus zR^3 \\ (3,2,1) &\in (z^{-1}\Lambda_1 \cap \Lambda_2 \cap \Lambda_3) \setminus zR^3. \end{aligned}$$

With this new matrix M', the images of Λ_i under the map $\Psi_{M'}$ become:

$$\begin{pmatrix} \Psi_{M'}(\Lambda_1) \\ \Psi_{M'}(\Lambda_2) \\ \Psi_{M'}(\Lambda_3) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & -1 & -1 & 0 & -1 & -1 & 0 & 0 & -1 \\ 0 & -1 & -1 & 0 & 0 & 0 & 0 & -1 & -1 & 0 & -1 & 0 \\ 0 & -1 & -1 & 0 & -1 & -1 & 0 & 0 & 0 & -1 & 0 & 0 \end{pmatrix}$$

This new membrane $[M']$ contains all the lattices in the min-convex hull of Λ_1, Λ_2, and Λ_3. The tropical convex hull of the three rows contains four other distinct

lattice points:

$$\begin{aligned}
\Psi_{M'}(\Lambda_1 \cap \Lambda_2) &= \Psi_{M'}(\Lambda_1) \oplus \Psi_{M'}(\Lambda_2), \\
\Psi_{M'}(\Lambda_1 \cap \Lambda_3) &= \Psi_{M'}(\Lambda_1) \oplus \Psi_{M'}(\Lambda_3), \\
\Psi_{M'}(\Lambda_2 \cap \Lambda_3) &= \Psi_{M'}(\Lambda_2) \oplus \Psi_{M'}(\Lambda_3), \\
\Psi_{M'}(\Lambda_1 \cap \Lambda_2 \cap \Lambda_3) &= \Psi_{M'}(\Lambda_1) \oplus \Psi_{M'}(\Lambda_2) \oplus \Psi_{M'}(\Lambda_3).
\end{aligned}$$

The simplicial complex $\text{minconv}(\Lambda_1, \Lambda_2, \Lambda_3)$ is shown on the right in Figure 4. □

5.3. **Computing max-convex hulls.** Algorithm 2 solves Computational Problem A in the min-convex case. Computing max-convex hulls reduces to computing min-convex hulls, as shown in Algorithm 3.

Input: matrices $M_1, M_2, \ldots, M_s \in \mathrm{GL}_d(K)$
Output: $\text{maxconv}(\Lambda_1, \ldots, \Lambda_s)$ in $\mathcal{B}_d$, where $\Lambda_i = \text{image}_R(M_i)$
Run Algorithm 2 with input matrices $M_1^{-T}, \ldots, M_s^{-T}$.
return $\text{minconv}(\Lambda_1^*, \ldots, \Lambda_s^*)$.

Algorithm 3: Max-convex hull in the Bruhat–Tits building $\mathcal{B}_d$.

The correctness of Algorithm 3 follows from Lemma 2, which implies that the simplicial complex structure of the max-convex hull of $\Lambda_1, \ldots, \Lambda_s$ is identical to the simplicial complex structure of the min-convex hull of $\Lambda_1^*, \ldots, \Lambda_s^*$. Our procedure exhibits a matrix of basis vectors for each lattice in $\text{minconv}(\Lambda_1^*, \ldots, \Lambda_s^*)$. We take the inverse transpose of that matrix to get a basis matrix for the corresponding lattice in $\text{maxconv}(\Lambda_1, \ldots, \Lambda_s)$.

There is a more straightforward way of computing max-convex hulls without using duality. Recall that membranes are max-convex. If we start with a finite set of lattices in a membrane $[M]$, the max-convex hull is a subcomplex of $[M]$, which is not apparent from Algorithm 3. Alternatively, Algorithm 4 computes the max-convex hull directly as a subcomplex of the membrane $[M]$. Its correctness follows from this proposition:

Proposition 25. *The max-convex hull in $\mathcal{B}_d$ of a finite set of lattices $\Lambda_1, \ldots, \Lambda_s$ in a membrane $[M]$ is the image under the nearest point map π_L of the max tropical convex hull of $\Psi_M(\Lambda_1), \ldots, \Psi_M(\Lambda_s)$ onto the tropical linear space L of the row space of M.*

Proof. Suppose $\Lambda, \Lambda' \in [M]$ are such that $\Psi_M(\Lambda) = \mathbf{a} \in L$ and $\Psi_M(\Lambda') = \mathbf{b} \in L$. Then $\Lambda = \langle z^{-a_1} f_1, \ldots, z^{-a_n} f_n \rangle$ and $\Lambda' = \langle z^{-b_1} f_1, \ldots, z^{-b_n} f_n \rangle$, so

$$\Lambda + \Lambda' = \langle z^{-max(a_1,b_1)} f_1, \ldots, z^{-max(a_n,b_n)} f_n \rangle.$$

Hence $\Lambda + \Lambda' \in [M]$, and $\Psi_M(\Lambda + \Lambda') = \pi_L(\max(\mathbf{a}, \mathbf{b}))$. The proposition follows directly from this. □

In the special case when the membrane M is an apartment, which is both max- and min- convex, the nearest point map is unnecessary, so Algorithm 4 reduces to computing the max-tropical convex hull in the tropical projective space.

Example 26. Let us compute the max-convex hull of the three lattices in Example 24 above. The max-tropical convex hull of the three rows of (12) contains 4 more

Input: matrices $M_1, M_2, \ldots, M_s \in \mathrm{GL}_d(K)$
Output: maxconv$(\Lambda_1, \ldots, \Lambda_s)$ in $\mathcal{B}_d$, where $\Lambda_i = \mathrm{image}_R(M_i)$
$M \leftarrow (M_1, \ldots, M_s)$.
$L \leftarrow$ tropical linear space of the row space of M.
for $i \leftarrow 1, 2, \ldots, s$ **do**
  $\Psi_M(\Lambda_i) \leftarrow$ tropical sum of the rows of $\mathrm{val}(M_i^{-1} \cdot M)$
$P \leftarrow$ max-tropical convex hull of $\Psi_M(\Lambda_1), \ldots, \Psi_M(\Lambda_s)$
Compute $\pi_L(P)$ using, for example, the Blue Rule or the Red Rule for each integer point in P.
return $\pi_L(P)$.

Algorithm 4: Max-convex hull in the Bruhat–Tits building $\mathcal{B}_d$.

points:

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

corresponding to the sums $\Lambda_1+\Lambda_2$, $\Lambda_1+\Lambda_3$, $\Lambda_2+\Lambda_3$, and $\Lambda_1+\Lambda_2+\Lambda_3$ respectively. However, all four lattices are equal to R^3, so their images under π_L must be the same point, and the max-convex hull contains just one more vertex than the original three. □

By duality again, Algorithm 4 can also be used to compute the min-convex hulls.

5.4. **Implementations.** We now come to question of how our convex hull algorithms can be used in practice, and what implementations are within reach. We largely focus on the operator "tconv" which is crucial in Algorithm 1, which in turn is called twice in Algorithm 2. Its output form (and hence also the form of the final output of the algorithm) were left deliberately vague, as there are several choices for how "tconv" can be realized. Firstly, there is a direct polyhedral approach for computing tropical convex hulls which is based on the following result from [7, Section 4]: The tropical convex hull of n points in $\mathbb{TP}^{s-1}$ arises as the polyhedral complex of bounded faces in an ordinary convex polyhedron defined by ns linear inequalities in $\mathbb{R}^{n+s}$. This method is implemented in `polymake` [11]. The details of this implementation together with extensive tests are the topic of [13]. Secondly, one can use the algebraic algorithm based on resolutions of monomial ideals which was described in [4]. A `Macaulay2`/`Maple` implementation is available from the third author. In the planar case, $s = 3$, specific techniques from computational geometry can be used to design alternative, faster algorithms; see [15].

In view of tropical polytope duality [7, Theorem 23], we can choose if we want to compute the tropical convex hull of n points in $\mathbb{TP}^{s-1}$ or of s points in $\mathbb{TP}^{n-1}$. If $s \leq 3$ then, due to the specialized algorithms mentioned above, it is easier to compute the tropical convex hull of n points in $\mathbb{TP}^{s-1}$. The output of both, the polyhedral and the algebraic algorithms, returns a tropical polytope P decomposed into cells as in (3).

Enumerating the lattices in Step 1 then requires to list all the lattice points in the ordinary polytopes corresponding to the types. In higher dimensions this can be an arduous task, due to the sheer size of the output. Hence, depending on the

application intended, it may be advisable to stick with the output of the previous stage as a compressed description of the set of lattices. From each type we can read off the matroid M_u which specifies the set of apartments (spanned by the columns of M) containing that type. In Example 3, these matroids M_u are the sets of pairs such as $\{af, bf, cf, df, ef, fg, fh\}$.

TABLE 1. Timings in seconds for computations with "tconv" in `polymake`. The parameters d and s indicate the size of the problem, that is, computing the min-convex hull of s lattices represented by $d \times d$-matrices. N is the number of samples tested, and the last four columns contain basic statistics.

d	s	N	mean	stddev	min	max
3	2	50	0.18	0.02	0.15	0.21
3	3	50	0.55	0.14	0.31	0.88
3	4	50	2.02	0.94	0.68	5.47
3	5	50	7.73	2.77	2.92	14.25
3	6	50	18.27	8.21	5.40	45.78
3	7	50	38.78	15.21	9.30	77.65
3	8	50	69.39	23.21	30.02	124.05
3	9	50	119.63	41.90	27.66	243.25
3	10	50	231.17	111.22	71.89	594.95
4	2	50	2.75	1.30	0.79	6.07
4	3	50	62.79	42.54	12.20	178.97
4	4	50	827.37	624.19	93.74	3017.19
4	5	18	5994.15	4986.38	648.14	21018.16
4	6	5	35823.43	21936.56	4846.15	67876.56
4	7	5	28266.78	15773.94	9193.69	55891.92

To give a sense of the running time of tropical convex hull code, in Table 1 we list a few timings of `polymake` computations. The samples were generated at random from $s \times sd$-matrices with integer entries ranging from 0 to 9. The algorithm uses the general polyhedral approach without the enumeration of lattice points. The individual timings vary quite a bit, and individual examples with smaller parameters may need more time than other examples with larger parameters. Nonetheless, the reader should get an idea. For more comprehensive tests we refer to [13]. Hardware: AMD 4200+X2, 4423 bogomips, 2GB main memory. Software implemented in `polymake` 2.3 on SuSE Linux 10.0.

6. Further Algorithms and Perspectives

We now consider Computational Problem B: Determine the intersection of s membranes. The input consists of matrices $M_1, \dots, M_s$, each having d linearly independent rows over $K = \mathbb{C}(\!(z)\!)$. Here M_i represents the membrane $[M_i] = [(f_{i1}, \dots, f_{id})]$, where f_{ij} is the jth column of the matrix M_i. The intersection $[M_1] \cap [M_2] \cap \cdots \cap [M_s]$ is a locally finite simplicial complex of dimension $\leq d-1$. It may be finite or infinite, depending on the input. We will compute this intersection as a tropical polytope over $(\mathbb{R} \cup \{\infty\}, \oplus, \odot)$.

Obviously, $[M_1] \cap [M_2] \cap \cdots \cap [M_s]$ is contained in the union $[M_1] \cup [M_2] \cup \cdots \cup [M_s]$, which in turn is contained in the membrane $[(M_1, M_2, \dots, M_s)]$. By Theorem 18,

this membrane is isomorphic, as a simplicial complex, to the standard triangulation of the tropicalization $L_p(M)$ of the row space of $M = (M_1, M_2, \ldots, M_s)$. In view of Theorem 14, we may regard $L_p(M)$ as a polytope in the compactified tropical projective space $\overline{\mathbb{TP}}^{sd-1}$.

Our computations take place inside this tropical linear space $L_p(M)$, which we represent as the tropical convex hull of the cocircuits $p(\sigma *)$ that are derived from the matrix M. The k-th column vector f_{ik} of the i-th input matrix M_i corresponds to the cocircuit $p(\sigma *)$ where σ is the $(d-1)$-subset of $\{1, 2, \ldots, sd\}$ which indexes all columns of M_i other than f_{ik} inside M. This special cocircuit is abbreviated by $C_{ik} := \mathrm{val}\big(\text{the } k\text{-th row } M_i^{-1} \cdot M\big)$. Consider the subpolytope of $L_p(M)$ spanned by the d special cocircuits arising from M_i:

$$L_p^M(M_i) \quad = \quad \mathrm{tconv}\{C_{i1}, \ldots, C_{id}\}.$$

This tropical polytope with its standard triangulation is isomorphic to the membrane $[M_i]$. Intersecting these subpolytopes $L_p(M_i)$ inside $L_p(M)$ solves Computational Problem B.

The intersections of arbitrary tropical polytopes are tropical polytopes again [7, Proposition 20]. Here, however, the situation is even easier since the subpolytope $L_p^M(M_i)$, as an ordinary polytopal complex, is a subcomplex of $L_p(M)$. We summarize our findings in Algorithm 5. Our remarks concerning the output of Algorithm 2 apply accordingly.

Input: Matrices $M_1, M_2, \ldots, M_s \in \mathrm{GL}_d(K)$
Output: Intersection $[M_1] \cap [M_2] \cap \cdots \cap [M_s]$ of membranes in $\mathcal{B}_d$
$M \leftarrow (M_1, M_2, \ldots, M_s)$
$C \leftarrow sd \times sd$-matrix of cocircuits of M
$L_p(M) \leftarrow \mathrm{tconv}\{c_{11}, \ldots, c_{ss}\}$
for $k \leftarrow 1, 2, \ldots, s$ **do**
  $L_p^M(M_i) \leftarrow \mathrm{tconv}\{c_{i1}, \ldots, c_{id}\}$
$I \leftarrow \emptyset$
foreach cell C in $L_p(M)$ **do**
  if $C \subseteq L_p^M(M_i)$ for all i **then**
    $I \leftarrow I \cup C$
return I

Algorithm 5: Intersection of membranes in the affine building $\mathcal{B}_d$

We now examine the special case of Computational Problem B where each input matrix M_i is square. Here our problem is to compute the intersection of s apartments in $\mathcal{B}_d$. Since apartments are both min- and max-convex, the intersection of apartments is also min- and max-convex. This establishes the connection between Computational Problem B and the classical notion of convexity in Remark 6. The set of all chambers which are fully contained in the intersection of apartments is convex in the sense of Remark 6. Note that (the vertex set of) every convex set of chambers within some apartment of $\mathcal{B}_d$ arises in this manner, namely as the output of Algorithm 5 for some square matrices $M_1, \ldots, M_s$. Identifying one of the apartments with $\mathbb{TP}^{d-1}$, we see that the result of this computation is a subset of $\mathbb{TP}^{d-1}$ which is both min-convex and max-convex. This implies that the intersection of apartments is an ordinary convex polytope of the special form (3).

Recent work of Alessandrini [2] suggests the following alternative method this computation, which more efficient than applying Algorithm 5 to square matrices. Our point of departure towards Alessandrini's method is the following question: *Given $M \in \mathrm{GL}_d(K)$, how can we decide whether the standard lattice R^d lies in the apartment $[M]$, i.e. whether R^d has an R-basis of the form $\{z^{a_1}f_1, z^{a_2}f_2, ..., z^{a_d}f_d\}$ for some integers $a_1, a_2, \ldots, a_d$?*

To answer this question, we compute the tropical $d \times d$-matrix

$$E(M) \quad := \quad \mathrm{val}(M) \odot \mathrm{val}(M^{-1}) \,. \tag{13}$$

Here $\odot$ means that the matrix product is evaluated in the min-plus algebra. Note that each diagonal entry of $E(M)$ is non-negative. The following lemma is easy to derive:

Lemma 27. *The following are equivalent for a matrix $M \in \mathrm{GL}_d(K)$:*

(a) *The standard lattice R^d lies in the apartment $[M]$.*
(b) *By scaling the columns of M with powers of z, we can get a matrix G in $R^{d\times d}$ whose constant term $G(0) \in \mathbb{C}^{d\times d}$ is invertible.*
(c) *Each entry $e_{ij}(M)$ of the matrix $E(M)$ is non-negative.*

We now change the question as follows. Let $u_1, \ldots, u_d$ be unknown integers. Under what condition on these integers is the scaled standard lattice $R\{z^{u_1}e_1, ..., z^{u_d}e_d\}$ in the apartment $[M]$? This question is equivalent to asking whether the standard lattice R^d lies in the apartment $[\,\mathrm{diag}(z^{-u}) \cdot M\,]$, where $\mathrm{diag}(z^{-u}) = \mathrm{diag}(z^{-u_1}, \ldots, z^{-u_d})$. By applying Lemma 27 to the matrix $\mathrm{diag}(z^{-u}) \cdot M$ in place of M, we obtain the following result.

Corollary 28. *The lattice $R\{z^{u_1}e_1, ..., z^{u_d}e_d\}$ lies in the apartment $[M]$ if and only if*

$$u_j - u_i \;\leq\; e_{ij}(M) \qquad \text{for } i, j = 1, 2, \ldots, d. \tag{14}$$

The linear inequalities (14) in the unknowns $u_1, \ldots, u_d$ defines a convex subset of $\mathbb{TP}^{d-1}$ which is both an ordinary polytope and a tropical polytope. Corollary 28 is essentially equivalent to Theorem 4.7 in [2]. Alessandrini refers to the polytope (14) as the *inversion domain* associated with the tropical matrix product in (13); see [2, Proposition 3.4]. We conclude that the intersection of the two apartments $[M]$ and $[\mathrm{diag}(1, \ldots, 1)]$ equals the standard triangulation of the inversion domain, which is specified by the inequalities (14).

We now present our second method, to be called *Alessandrini's Algorithm*, for Computational Problem B in the special case of apartments. The input consists of s invertible matrices $M_1, M_2, \ldots, M_s$ over K, and the output is the intersection $[M_1] \cap \cdots \cap [M_s]$ of apartments. After multiplying each matrix on the left by M_1^{-1}, we may assume that M_1 is the identity matrix. Then the desired intersection is the standard triangulation of the polytope specified by the inequalities (14) where M runs over $\{M_2, \ldots, M_k\}$. Alessandrini's Algorithm is summarized by the following refinement of Proposition 20.

Theorem 29. *The intersection of apartments $[M_1] \cap \cdots \cap [M_s]$ in the Bruhat–Tits building $\mathcal{B}_d$ is the standard triangulation of a polytope of the form (3), namely, the polytope*

$$\left\{\, u \in \mathbb{TP}^{d-1} \,:\, u_j - u_i \leq e_{ij}(M_k) \text{ for } i, j = 1, \ldots, d \text{ and } k = 2, \ldots, s \,\right\} .$$

Conclusion

We have demonstrated that tropical convexity is a useful tool for computations with affine buildings. Given the ubiquitous appearance of affine buildings in mathematics, we are optimistic that our approach can be of interest for a wide range of applications. Such applications may arise in fields as diverse as geometric topology [2], number theory [10, 21], algebraic geometry [16, 17], representation theory [12], harmonic analysis [19], and differential equations [6]. Experts in combinatorial representation theory may find it interesting to generalize our constructions and algorithms to affine buildings of other types. This will require to investigate, for instance, the B_n-analogs of tropical polytopes.

References

[1] Peter Abramenko and Kenneth S. Brown, *Approaches to Buildings*, Springer-Verlag, New York, 2007.

[2] Daniele Alessandrini, *Tropicalization of group representations*, `arXiv:math.GT/0703608`.

[3] Federico Ardila, *Subdominant matroid ultrametrics*, Annals of Combinatorics **8** (2004) 379–389.

[4] Florian Block and Josephine Yu, *Tropical convexity via cellular resolutions*, J. Algebraic Combin. **24** (2006), no. 1, 103–114.

[5] François Bruhat and Jacques Tits, *BN-paires de type affine et données radicielles*, C. R. Acad. Sci. Paris Sér. A-B **263** (1966), A598–A601.

[6] Eduardo Corel, *Moser-reduction of lattices for a linear connection*, preprint, 2007, `www.math.jussieu.fr/~corel/publications/publi-list.html`.

[7] Mike Develin and Bernd Sturmfels, *Tropical convexity*, Doc. Math. **9** (2004), 1–27 (electronic).

[8] Andreas Dress and Werner Terhalle, *A combinatorial approach to* $\mathfrak{p}$*-adic geometry*, Geom. Dedicata **46** (1993), no. 2, 127–148.

[9] ———, *The tree of life and other affine buildings*, Proceedings of the International Congress of Mathematicians, Vol. III (Berlin, 1998), Extra Vol. III, 1998, pp. 565–574 (electronic).

[10] Gerd Faltings, *Toroidal resolutions for some matrix singularities*, Moduli of abelian varieties (Texel Island, 1999), Progr. Math., vol. 195, Birkhäuser, Basel, 2001, pp. 157–184.

[11] Ewgenij Gawrilow and Michael Joswig, *polymake: a framework for analyzing convex polytopes*, Polytopes—combinatorics and computation (Oberwolfach, 1997), DMV Sem., vol. 29, Birkhäuser, Basel, 2000, pp. 43–73.

[12] Ulrich Görtz, *Alcove walks and nearby cycles on affine flag manifolds*, Journal of Algebraic Combinatorics, **26** (2007) 415-430.

[13] Sven Herrmann, Michael Joswig, and Marc E. Pfetsch, *Computing the bounded subcomplex of an unbounded polyhedron*, in preparation.

[14] Petra Hitzelberger, *A convexity theorem for affine buildings*, `arXiv:math.MG/0701094`.

[15] Michael Joswig, *Tropical halfspaces*, Combinatorial and computational geometry, Math. Sci. Res. Inst. Publ., vol. 52, Cambridge Univ. Press, Cambridge, 2005, pp. 409–431.

[16] Mikhail M. Kapranov, *Chow quotients of Grassmannians. I*, I. M. Gel′fand Seminar, Adv. Soviet Math., vol. 16, Amer. Math. Soc., Providence, RI, 1993, pp. 29–110.

[17] Sean Keel and Jenia Tevelev, *Geometry of Chow quotients of Grassmannians*, Duke Math. J. **134** (2006), no. 2, 259–311.

[18] G. A. Mustafin, *Non-Archimedean uniformization*, Mat. Sb. (N.S.) **34** (1978), no. 2, 187–214.

[19] James Parkinson, *Spherical harmonic analysis on affine buildings*, Mathematische Zeitschrift **253** (2006), no. 3, 571–606.

[20] Jürgen Richter-Gebert, Bernd Sturmfels and Thorsten Theobald, *First steps in tropical geometry*, in "Idempotent Mathematics and Mathematical Physics", Proceedings Vienna 2003, (editors G.L. Litvinov, V.P. Maslov), American Math. Society, Contemporary Mathematics **377** (2005) 289-317.

[21] Alison Setyadi, *Distance in the Affine Buildings of* SL_n *and* Sp_n, `arXiv:math.NT/0511556`

[22] David Speyer, *Tropical linear spaces*, `arXiv:math.CO/0410455`.

[23] David Speyer and Bernd Sturmfels, *The tropical Grassmannian*, Adv. in Geometry **4** (2004) 389-411.

[24] Jacques Tits, *Buildings of spherical type and finite BN-pairs*, Springer-Verlag, Berlin, 1974, Lecture Notes in Mathematics, Vol. 386.
[25] Annette Werner, *Compactification of the Bruhat-Tits building of PGL by lattices of smaller rank*, Doc. Math. **6** (2001), 315–341 (electronic).
[26] ———, *Compactification of the Bruhat-Tits building of PGL by seminorms*, Math. Z. **248** (2004), no. 3, 511–526.
[27] Josephine Yu and Debbie S. Yuster, *Representing tropical linear spaces by circuits*, `arXiv:math.CO/0611579`.

Michael Joswig,
FB Mathematik, AG 7,
TU Darmstadt, 64289 Darmstadt,
Germany
Email: joswig@mathematik.tu-darmstadt.de

Bernd Sturmfels
Department of Mathematics,
University of California at Berkeley,
Berkeley CA 94720, USA
Email: bernd@math.berkeley.edu

Josephine Yu
Department of Mathematics,
M.I.T.,
Cambridge MA 02139, USA
Email: jyu@math.mit.edu

ALBANIAN JOURNAL
OF MATHEMATICS
Volume 1, Number 4, Pages 213–221
ISSN 1930-1235: (2007)

DISCRETE SERIES REPRESENTATIONS OF p-ADIC GROUPS ASSOCIATED TO SYMMETRIC SPACES

JEFFREY HAKIM

1. Introduction

The purpose of this paper is study the natural symmetric space analogues of various notions related to discrete series representations of a p-adic group such as Schur's orthogonality relations and formal degrees.

We study representations of the group $G = \mathbf{G}(F)$ of F-rational points of a connected, reductive F-group $\mathbf{G}$, where F is a finite extension of a field $\mathbb{Q}_p$ of p-adic numbers for some odd prime p.

The representations of interest are associated to a symmetric space $H\backslash G$, where $H = \mathbf{H}(F)$ and $\mathbf{H}$ is the group of fixed points of some F-automorphism θ of $\mathbf{G}$ of order two.

To be more precise, we are interested in irreducible admissible complex representations (π, V) of G that are *H-distinguished* in the sense that there exists a nonzero linear form $\tilde{\lambda} : V \to \mathbb{C}$ that is H-fixed or, in other words,

$$\langle \pi(h)v, \tilde{\lambda} \rangle = \langle v, \tilde{\lambda} \rangle,$$

for all $h \in H$ and $v \in V$. From now on, assume that such a representation (π, V) has been fixed. Note that H-distinction[1] implies that the restriction of the central quasi-character of π to Z_H is trivial.

The latter linear forms, together with 0, comprise the space $\mathrm{Hom}_H(\pi, 1)$ and Frobenius Reciprocity maps this space isomorphically onto the space

$$\mathrm{Hom}_G(\pi, C^\infty(H\backslash G)),$$

where $C^\infty(H\backslash G)$ is the space of smooth complex-valued functions on $H\backslash G$ viewed as a G-module with respect to right translations by G. So π is H-distinguished precisely when it has a G-invariant embedding in $C^\infty(H\backslash G)$. In this sense, the H-distinguished representations are precisely the representations that contribute to the harmonic analysis on $H\backslash G$.

For convenience, we will make some simplifying assumptions that are generally satisfied in applications. Let $(\tilde{\pi}, \widetilde{V})$ be the contragredient of (π, V). We assume that $\mathrm{Hom}_H(\tilde{\pi}, 1)$, in addition to $\mathrm{Hom}_H(\pi, 1)$, is nonzero and, furthermore, we assume that both of the latter spaces are finite-dimensional.

Let $\mathbf{Z}$ be the center of $\mathbf{G}$ and let $\mathbf{Z_H} = \mathbf{Z} \cap \mathbf{H}$ and let $Z = \mathbf{Z}(F)$ and $Z_H = \mathbf{Z_H}(F)$. We fix a Haar measure on H/Z_H for use in our integrations over the latter quotient.

[1] At the suggestion of Hervé Jacquet, we refer to the property of being H-distinguished as "H-distinction," rather than "H-distinguishedness."

Recall that (π, V) is said to be a *discrete series representation* if its central quasi-character is unitary and the absolute value of every matrix coefficient of π lies in $L^2(G/Z)$.

Definition. *If the restriction to H of every matrix coefficient of π lies in $L^1(H/Z_H)$ then π is said to be* θ**-discrete**. We now give a symmetric space analogue of this notion.

When π is θ-discrete, we may define the pairing

$$\langle v, \tilde{v} \rangle_\theta = \int_{H/Z_H} \langle \pi(h)v, \tilde{v} \rangle \, dh$$

for $v \in V$ and $\tilde{v} \in \widetilde{V}$. To appreciate the above terminology, one should consider the so-called "group case." In this case, $\mathbf{G} = \mathbf{G}_1 \times \mathbf{G}_1$, for some connected reductive F-group $\mathbf{G}_1$ and let $\theta(g_1, g_2) = (g_2, g_1)$. Then if (π, V) is H-distinguished it must have the form $(\pi_1 \times \tilde{\pi}_1, V_1 \otimes \widetilde{V}_1)$. The contragredient of (π, V) is then $(\tilde{\pi}_1 \times \pi_1, \widetilde{V}_1 \otimes V_1)$ and the invariant pairing on $V \times \widetilde{V}$ is just

$$\langle v \otimes \tilde{v}, \tilde{u} \otimes u \rangle = \langle v, \tilde{u} \rangle \, \langle u, \tilde{v} \rangle.$$

Thus

$$\langle v \otimes \tilde{v}, \tilde{u} \otimes u \rangle_\theta = \int_{G_1/Z(G_1)} \langle \pi_1(g)v, \tilde{u} \rangle \, \langle u, \tilde{\pi}_1(g)\tilde{v} \rangle \, dg,$$

where $Z(G_1)$ is the center of G_1. These integrals occur in Schur's orthogonality relations when π_1 is a discrete series representation.

Definition. *If π is θ-discrete and $v \in V$ and $\tilde{v} \in \widetilde{V}$ then the function $f^\theta_{v \otimes \tilde{v}} \in C^\infty(H \backslash G)$ defined by*

$$f^\theta_{v \otimes \tilde{v}}(g) = \langle \pi(g)v, \tilde{v} \rangle_\theta$$

is called a θ**-matrix coefficient** *of π.*

Definition. *If π is θ-discrete and every θ-matrix coefficient of π is supported in a compact subset of $ZH \backslash G$ then we say that π is* θ**-supercuspidal**.

Definition. (Kato and Takano [KT]) *If the function*

$$f^\theta_{v \otimes \tilde{\lambda}}(g) = \langle \pi(g)v, \tilde{\lambda} \rangle$$

is supported in a compact subset of $ZH \backslash G$, for all $v \in V$ and all $\tilde{\lambda} \in \mathrm{Hom}_H(\pi, 1)$ then we say that π is H**-relatively cuspidal**.

If π is θ-discrete and $\tilde{v} \in \widetilde{V}$ then there is an associated invariant linear form $\tilde{\lambda}_{\tilde{v}} \in \mathrm{Hom}(\pi, 1)$ by

$$\langle v, \tilde{\lambda}_{\tilde{v}} \rangle = \langle v, \tilde{v} \rangle_\theta.$$

The following lemma follows easily from the latter fact:

Lemma. *If π is supercuspidal or H-relatively cuspidal then it is θ-supercuspidal.*

2. Formal degrees and orthogonality relations

If π is a discrete series representation then the Schur orthogonality relations hold and they say that there exists a nonzero constant $d(\pi)$ (depending on the choice of a Haar measure on G/Z) such that

$$\int_{G/Z} \langle \pi(g)v, \tilde{u}\rangle \ \langle u, \tilde{\pi}(g)\tilde{v}\rangle \ dg = d(\pi)^{-1}\langle v, \tilde{v}\rangle \ \langle u, \tilde{u}\rangle,$$

for all $u, v \in V$ and $\tilde{u}, \tilde{v} \in \widetilde{V}$. The constant $d(\pi)$ is called *the formal degree of* π (with respect to the given measure on G/Z).

It is well known that if π is a supercuspidal representation that is compactly induced from an irreducible representation ρ of an open compact-mod-center subgroup K of G then $d(\pi)$ is quotient of the degree of ρ and the measure of the image of K in G/Z. One can find a proof of this fact in [M1]. We generalize both the statement of this result and the proof later in this paper.

2.1. **The multiplicity one case.** In this section, we consider symmetric space generalizations of the formal degree and Schur's orthogonality relations. We make the simplifying assumption that the spaces $\mathrm{Hom}_H(\tilde{\pi}, 1)$ and $\mathrm{Hom}_H(\pi, 1)$ have dimension one and we fix nonzero linear forms $\lambda \in \mathrm{Hom}_H(\tilde{\pi}, 1)$ and $\tilde{\lambda} \in \mathrm{Hom}_H(\pi, 1)$.

Let ω denote the central character of the discrete series representation (π, V). Let $C^\infty(G, \omega)$ denote the space of smooth complex-valued functions f on G such that

$$f(zg) = \omega(z)^{-1} f(g),$$

for all $z \in Z$ and $g \in G$. For such f, we may define a vector $\pi(f)\lambda \in V$ by the relation

$$\langle \pi(f)\lambda, \tilde{v}\rangle = \int_{G/Z} f(g) \ \langle \lambda, \tilde{\pi}(g)^{-1}\tilde{v}\rangle \ dg,$$

for all $\tilde{v} \in \widetilde{V}$. We also define

$$\Theta_{\lambda\otimes\tilde{\lambda}}(f) = \langle \pi(f)\lambda, \tilde{\lambda}\rangle.$$

This is a linear functional on $C^\infty(G, \omega)$ and it is a generalized matrix coefficient distribution. (In fact, it is a distribution on the ℓ-sheaf $C^\infty(G, \omega)$, in the sense of [BZ].) If we take the test function f to be a matrix coefficient

$$f_{\tilde{v}\otimes v}(g) = \langle v, \tilde{\pi}(g)\tilde{v}\rangle$$

of $\tilde{\pi}$ then we have the following extension of Schur's orthogonality relations:

Lemma. ([H]) $\quad \Theta_{\lambda\otimes\tilde{\lambda}}(f_{\tilde{v},v}) = d(\pi)^{-1}\langle \lambda, \tilde{v}\rangle \ \langle v, \tilde{\lambda}\rangle.$

Proof. Fix a compact open subgroup K of G that fixes v and $\tilde{v}$. Let e_K be the convolution idempotent associated to K. Let $\lambda_K = \pi(e_K)\lambda \in V$ and $\tilde{\lambda}_K = \tilde{\pi}(e_K)\tilde{\lambda} \in \widetilde{V}$. Then

$$\begin{aligned}
\Theta_{\lambda\otimes\tilde{\lambda}}(f_{\tilde{v},v}) &= \int_{G/Z} \langle \pi(g)\lambda_K, \tilde{\lambda}_K\rangle \ \langle v, \tilde{\pi}(g)\tilde{v}\rangle \ dg \\
&= d(\pi)^{-1}\langle \lambda_K, \tilde{v}\rangle \ \langle v, \tilde{\lambda}_K\rangle \\
&= d(\pi)^{-1}\langle \lambda, \tilde{v}\rangle \ \langle v, \tilde{\lambda}\rangle.
\end{aligned}$$

□

Proposition. *There exists a unique nonzero constant $d_\theta(\pi)$ (depending on the choice of measure on H/Z_H used to define $\langle\ ,\ \rangle_\theta$ and on the choices of λ and $\tilde{\lambda}$) such that*

$$\langle v, \tilde{v}\rangle_\theta = d_\theta(\pi)^{-1}\langle v, \tilde{\lambda}\rangle\ \langle \lambda, \tilde{v}\rangle$$

for all θ-discrete representations (π, V) and all $v \in V$ and $\tilde{v} \in \widetilde{V}$.

Proof. Consider $\langle v, \tilde{v}\rangle_\theta$ as $\tilde{v}$ is fixed and v varies. This defines an element of $\mathrm{Hom}_H(\pi, 1)$ and hence a multiple of $\tilde{\lambda}$. Thus there exists a complex number $\gamma(\tilde{v})$ such that $\langle v, \tilde{v}\rangle_\theta = \gamma(\tilde{v})\ \langle v, \tilde{\lambda}\rangle$. Now consider $\gamma(\tilde{v})$ as $\tilde{v}$ varies. This must be a multiple of λ. Therefore, there must be a constant c such that

$$\langle v, \tilde{v}\rangle_\theta = c\ \langle v, \tilde{\lambda}\rangle\ \langle \lambda, \tilde{v}\rangle.$$

Now choose v and $\tilde{v}$ so that $\langle v, \tilde{\lambda}\rangle$ and $\langle \lambda, \tilde{v}\rangle$ are nonzero. From the previous lemma, we have

$$\Theta_{\lambda\otimes\tilde{\lambda}}(f_{\tilde{v},v}) = d(\pi)^{-1}\langle v, \tilde{\lambda}\rangle\ \langle \lambda, \tilde{v}\rangle$$

and thus $\Theta_{\lambda\otimes\tilde{\lambda}}(f_{\tilde{v},v})$ is nonzero. This implies that

$$\int_{H/Z_H} f_{\tilde{v},v}(hg)\ dh$$

is not identically zero for all $g \in G$. But the latter integral is just $\langle v, \tilde{\pi}(g)\tilde{v}\rangle$. This shows that the pairing $\langle\ ,\ \rangle_\theta$ is not identically zero. It follows that c is nonzero. Taking $d_\theta(\pi) = c^{-1}$ completes the proof. □

2.2. **The group case.** Consider the group $G \times G$ with the involution $\theta(a, b) = (b, a)$. Fix an irreducible, smooth representation π of G and let $\Pi = \pi \times \tilde{\pi}$ and $\widetilde{\Pi} = \tilde{\pi} \times \pi$. Let $\lambda : \widetilde{V} \times V \to \mathbb{C}$ and $\tilde{\lambda} : V \times \widetilde{V} \to \mathbb{C}$ be the obvious canonical pairings. Then

$$\langle u \otimes \tilde{u}, \tilde{v} \otimes v\rangle_\theta = \int_{G/Z} \langle \pi(x)u, \tilde{v}\rangle\ \langle v, \tilde{\pi}(x)\tilde{u}\rangle\ dx^*.$$

In addition,

$$\begin{aligned} \langle u \otimes \tilde{u}, \tilde{\lambda}\rangle &= \langle u, \tilde{u}\rangle, \\ \langle \lambda, \tilde{v} \otimes v\rangle &= \langle v, \tilde{v}\rangle. \end{aligned}$$

Therefore,

$$d_\theta(\pi \times \tilde{\pi}) = d(\pi).$$

2.3. **The finite multiplicity case.** Suppose that $\mathrm{Hom}_H(\tilde{\pi}, 1)$ has finite dimension n_1 and is spanned by elements $\lambda_1, \ldots, \lambda_{n_1}$. Suppose that $\mathrm{Hom}_H(\pi, 1)$ has finite dimension n_2 and is spanned by elements $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_{n_2}$. For fixed $\tilde{v} \in \widetilde{V}$, $\langle v, \tilde{v}\rangle_\theta$ defines an element of $\mathrm{Hom}_H(\pi, 1)$ and thus there exist numbers $c_j(\tilde{v})$ such that

$$\langle v, \tilde{v}\rangle_\theta = \sum_{j=1}^{n_2} c_j(\tilde{v})\langle v, \tilde{\lambda}_j\rangle.$$

On the other hand, for each j, it is easy to see that $c_j \in \mathrm{Hom}_H(\tilde{\pi}, 1)$ and so there exist numbers $D_\theta(\pi)_{ij}$ such that

$$c_j(\tilde{v}) = \sum_{i=1}^{n_1} D_\theta(\pi)_{ij}\langle \lambda_i, \tilde{v}\rangle.$$

Hence we have

$$\langle v, \tilde{v}\rangle_\theta = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} D_\theta(\pi)_{ij}\langle \lambda_i, \tilde{v}\rangle\langle v, \tilde{\lambda}_j\rangle.$$

2.4. **Induced supercuspidal representations.** Let K be a θ-stable subgroup of G that contains Z and is such that the quotient K/Z is compact. Let $K_H = K \cap H$. Let (ρ, W) be an irreducible (finite-dimensional) complex representation of K with unitary central character and let $(\tilde{\rho}, \widetilde{W})$ denote the contragredient.

The representations (π, V) and $(\tilde{\pi}, \widetilde{V})$ obtained by compactly-supported induction from (ρ, W) and $(\tilde{\rho}, \widetilde{W})$ are irreducible supercuspidal representations. We use the standard pairing $\langle\ ,\ \rangle$ on $W \times \widetilde{W}$ and use this to define an invariant pairing on $V \times \widetilde{V}$ as follows:

$$\langle v, \tilde{v}\rangle = \sum_{Kg\in K\backslash G} \langle v(g), \tilde{v}(g)\rangle.$$

Let $\{e_i\}$ and $\{\tilde{e}_j\}$ be bases of the spaces of K_H-fixed vectors in W and $\widetilde{W}$, respectively. Define

$$\langle w, \tilde{w}\rangle_\theta = \int_{K_H/Z_H} \langle \rho(h)w, \tilde{w}\rangle\ dh,$$

where we choose the Haar measure on K_H/Z_H so that K_H/Z_H has volume one. Given $w \in W$ and $\tilde{w} \in \widetilde{W}$, there exist unique coefficients w_i and $\tilde{w}_j$ such that

$$\int_{K_H/Z_H} \rho(h)w\ dh = \sum_i w_i e_i, \qquad \int_{K_H/Z_H} \tilde{\rho}(h)\tilde{w}\ dh = \sum_j \tilde{w}_j e_j.$$

We have

$$\langle w, \tilde{w}\rangle_\theta = \sum_{i,j} w_i \tilde{w}_j D_\theta(\rho)_{ij},$$

where

$$D_\theta(\rho)_{ij} = \langle e_i, \tilde{e}_j\rangle.$$

Given $w \in W$, there is an associated element $v_w \in V$ that is defined by $v_w(k) = \rho(k)w$, for all $k \in K$, and $v_w|(G - K) \equiv 0$. This is a K-equivariant embedding of ρ in π. By Frobenius reciprocity, ρ occurs in π with multiplicity one.

Similarly, we associate to $\tilde{w} \in \widetilde{W}$ an element $\tilde{v}_{\tilde{w}} \in \widetilde{V}$. The map $(w, \tilde{w}) \mapsto (v_w, \tilde{v}_{\tilde{w}})$ is an isometric embedding in the sense that $\langle w, \tilde{w}\rangle = \langle v_w, \tilde{v}_{\tilde{w}}\rangle$. We now compute $\langle v_w, \tilde{v}_{\tilde{w}}\rangle_\theta$. We use the measure on H/Z_H given by the integral formula

$$\int_{H/Z_H} f(h)\ dh = \sum_{h\in K_H\backslash H} \int_{K_H/Z_H} f(kh)\ dk.$$

We have

$$\begin{aligned}
\langle v_w, \tilde{v}_{\tilde{w}}\rangle_\theta &= \sum_{h\in K_H\backslash H} \int_{K_H/Z_H} \langle \pi(kh)v_w, \tilde{v}_{\tilde{w}}\rangle\ dk \\
&= \int_{K_H/Z_H} \langle \pi(k)v_w, \tilde{v}_{\tilde{w}}\rangle\ dk \\
&= \int_{K_H/Z_H} \langle \rho(k)w, \tilde{w}\rangle\ dk \\
&= \langle w, \tilde{w}\rangle_\theta.
\end{aligned}$$

Therefore, in the notation of the previous section,

$$\sum_{i,j} v_i \ \tilde{v}_j \ D_\theta(\pi)_{ij} = \sum_{ij} w_i \ \tilde{w}_j \ D_\theta(\rho)_{ij},$$

where $v_i = \langle v_w, \tilde{\lambda}_j \rangle$ and $\tilde{v}_j = \langle \lambda_i, \tilde{v}_{\tilde{w}} \rangle$.

To proceed further, we now exploit the fact that all of representations are necessarily unitarizable. More precisely, we let $\overline{W}$ be the set W together with its additive structure, but with scalar multiplication defined by $c \cdot w = \bar{c}w$. Letting $\bar{\rho}(k) = \rho(k)$, we obtain a representation $(\bar{\rho}, \overline{W})$ of K. Let $(\ ,\ )$ be an invariant non-degenerate hermitian form on W that realizes ρ as a unitary representation. Then $v \mapsto \tilde{v} = (\ , v)$ defines an isomorphism of $(\bar{\rho}, \overline{W})$ with $(\tilde{\rho}, \widetilde{W})$.

We take $\{e_i\}$ to be an orthonormal basis of the space of K_H-fixed vectors in W. Then $\{\tilde{e}_i\}$ is a dual basis of the K_H-fixed vectors in $\widetilde{W}$. In particular, the spaces of K_H-fixed vectors in W and $\widetilde{W}$ have the same dimension. The matrix $D_\theta(\rho)$ is now an identity matrix. The only invariant of the matrix is its rank (which is its trace) and taking $w = e_1 + \cdots + e_r$ we have

$$\langle w, \tilde{w} \rangle_\theta = \operatorname{trace}(D_\theta(\rho)) = r = \dim W^{K_H},$$

where W^{K_H} is the space of K_H-fixed vectors in W.

Now each basis element e_i gives an element $\tilde{e}_i \in \widetilde{W}$ and this yields an element $\tilde{v}_{\tilde{e}_i} \in \widetilde{V}$. We now define $\tilde{\lambda}_i \in \operatorname{Hom}_H(\pi, 1)$ by

$$\langle v, \tilde{\lambda}_i \rangle = \langle v, \tilde{v}_{\tilde{e}_i} \rangle_\theta.$$

This linear form must be nonzero since

$$\langle v_{e_i}, \tilde{\lambda}_i \rangle = \langle v_{e_i}, \tilde{v}_{\tilde{e}_i} \rangle_\theta = \langle e_i, \tilde{e}_i \rangle_\theta = 1.$$

Define invariant linear forms λ_j similarly on $\widetilde{V}$. Suppose that the latter linear forms span the H-invariant linear forms in V and $\widetilde{V}$. Then if $i, j \in \{1, \ldots, r\}$ then

$$D_\theta(\pi)_{ij} = \langle v_{e_j}, \tilde{v}_{\tilde{e}_i} \rangle_\theta = \delta_{ij}.$$

The space V is a space of functions on G. It has a natural decomposition

$$V = \bigoplus_{KgH \in K \backslash G / H} V_{KgH},$$

where V_{KgH} is the space of functions in V that have support contained in the double coset KgH. Let V^* be the space of linear forms on V and let $(V^*)^H$ be the subspace of H-fixed linear forms. Defining V^*_{KgH} and $(V^*_{KgH})^H$ similarly, we have

$$(V^*)^H = \bigoplus_{KgH \in K \backslash G / H} (V^*_{KgH})^H.$$

We may view the elements of $(V^*_{KH})^H$ as the *primary* elements of $(V^*)^H$. It is easy to see that the H- invariant linear forms $\tilde{\lambda}_i$ constructed above from elements of $\widetilde{W}$ are primary elements, in the latter sense, and, in fact, all primary elements are obtained in this manner. Indeed, this follows from the fact that $(V^*_{KgH})^H \cong \widetilde{W}^{K \cap gHg^{-1}}$. (See [HMa1] and [HMa2].) For examples of non-primary invariant linear forms see [HM].

Proposition. *Let (π, V) be an H-distinguished supercuspidal representation that is induced from an open compact-mod-center θ-stable subgroup K. Then every H-invariant linear form $\tilde{\lambda}$ on V has the form $\tilde{\lambda}_{\tilde{v}}$ for some $\tilde{v} \in \widetilde{V}$. Therefore, π must be H-relatively cuspidal.*

Proof. We may as well assume π has a unitary central character, since there is no harm in replacing π by a twist by a quasi-character. Next, we may as well assume that $\tilde{\lambda}$ is supported in a single double coset KgH. Then λ is associated to some element $\tilde{w} \in \widetilde{W}^{K \cap gHg^{-1}}$. Define $\tilde{v}$ by $\tilde{v}(kg) = \tilde{\rho}(k)\tilde{w}$, for all $k \in K$ and $\tilde{v}|(G - Kg) \equiv 0$. Then $\tilde{\lambda}_{\tilde{v}}$ is the element of $(V^*_{KgH})^H$ associated to $\tilde{w}$. Our claim follows. □

3. Orbital integrals

3.1. **The Harish-Chandra/Rader/Silberger formula.** Fix an elliptic Cartan subgroup Γ of G, that is, a Cartan subgroup such that Γ/Z is compact. Suppose f is a smooth function on G with compact-mod-center support. Define

$$F_f(\gamma) = |D(\gamma)|^{1/2} \int_{G/Z} f(g\gamma g^{-1})\ dg,$$

for $\gamma \in \Gamma' = \Gamma \cap G'$. Then F_f is a smooth function on Γ'.

We now recall the derivation of the supercuspidal case of an integral formula proved by Harish-Chandra [HC] and generalized by Rader and Silberger [RS].

Assume π is irreducible supercuspidal and $\gamma \in \Gamma'$ and $f(g) = \langle \pi(g)v, \tilde{v} \rangle$.

$$\begin{aligned}
\int_{G/Z} f(g\gamma g^{-1})\ dg &= \int_{G/Z} \langle \pi(g\gamma g^{-1})v, \tilde{v} \rangle\ dg \\
&= \int_{G/Z} \sum_i \langle \pi(g\gamma g^{-1})v, \tilde{\pi}(g\gamma)\tilde{e}_i \rangle\ \langle \pi(g\gamma)e_i, \tilde{v} \rangle\ dg \\
&= \int_{G/Z} \sum_i \langle \pi(g^{-1})v, \tilde{e}_i \rangle\ \langle \pi(\gamma)e_i, \tilde{\pi}(g^{-1})\tilde{v} \rangle\ dg \\
&= \sum_i \int_{G/Z} \langle \pi(g^{-1})v, \tilde{e}_i \rangle\ \langle \pi(\gamma)e_i, \tilde{\pi}(g^{-1})\tilde{v} \rangle\ dg \\
&= \sum_i \int_{G/Z} \langle \pi(g)v, \tilde{e}_i \rangle\ \langle \pi(\gamma)e_i, \tilde{\pi}(g)\tilde{v} \rangle\ dg \\
&= d(\pi)^{-1} \langle v, \tilde{v} \rangle \sum_i \langle \pi(\gamma)e_i, \tilde{e}_i \rangle \\
&= d(\pi)^{-1} \langle v, \tilde{v} \rangle\ \Theta_\pi(\gamma),
\end{aligned}$$

where Θ_π is the character of π.

It follows that if π is irreducible supercuspidal then

$$F_f(\gamma) = d(\pi)^{-1} f(1)\ |D(\gamma)|^{1/2}\ \Theta_\pi(\gamma)$$

for all $\gamma \in \Gamma'$ and all f in the vector space $\mathcal{A}(\pi)$ spanned by the matrix coefficients of π. This formula is an instance of the classic principle of duality between orbital integrals and characters.

3.2. **The symmetric space generalization.** The group $H \times H$ acts on G by $(h_1, h_2) \cdot g = h_1 g h_2^{-1}$. Let $(H \times H)_g$ be the isotropy group of g. Then, given a function f defined at least on HgH, one can consider the integral

$$\int_{(H\times H)/(H\times H)_g} f(h_1 g h_2^{-1})\ d(h_1, h_2),$$

assuming that it is possible to choose a nonzero invariant measure for the integration. For example, if $f(g) = \langle \pi(g)v, \tilde{v}\rangle$ and $g = 1$ then the orbital integral is just $\langle v, \tilde{v}\rangle$.

We assume that γ is an element of G such that $(H \times H)_\gamma$ is compact modulo $Z_H \times Z_H$. This is the case, for example, if γ is θ-*elliptic-regular* in the sense that $\gamma\theta(\gamma)^{-1}$ is elliptic regular.

Let (π, V) be a θ-discrete representation and assume that $\mathrm{Hom}_H(\tilde{,}1)$ and $\mathrm{Hom}_H(\pi, 1)$ are 1-dimensional and are spanned by nonzero elements λ and $\tilde{\lambda}$, respectively. Let Θ be the smooth function on the θ-regular set that represents the distribution $\Theta_{\lambda,\tilde{\lambda}}$. (See [RR] for more details on this terminology.)

Let

$$f(g) = \langle \pi(g)v, \tilde{v}\rangle$$

be a matrix coefficient for π. The un-normalized orbital integral of f at a θ-elliptic-regular element γ is

$$\Phi_f(\gamma) = \int_{(H/Z_H)^2} f(h_1\gamma h_2)\ dh_1\ dh_2.$$

Proposition. $\Phi_f(\gamma) = d_\theta^{-2}(\pi)\ \langle \lambda, \tilde{v}\rangle\ \langle v, \tilde{\lambda}\rangle\ \Theta(\gamma).$

Proof. We have

$$\begin{aligned}
\Phi_f(\gamma) &= \int_{(H/Z_H)^2} \langle \pi(h_1\gamma h_2)v, \tilde{v}\rangle\ dh_1\ dh_2 \\
&= \int_{H/Z_H} \langle \pi(\gamma h)v, \tilde{v}\rangle_\theta\ dh \\
&= d_\theta^{-1}(\pi)\ \langle \lambda, \tilde{v}\rangle \int_{G^\theta/Z^\theta} \langle \pi(\gamma h)v, \tilde{\lambda}\rangle\ dh.
\end{aligned}$$

We now invoke the key lemma of Rader and Rallis [RR] in their proof of smoothness on the θ-regular set of spherical characters. It says that we can choose a compact open subgroup K_H of H such that

$$\int_{K_H} \pi(k\gamma^{-1})\tilde{\lambda}\ dk$$

lies in $\widetilde{V}$. Call this vector $\tilde{v}_\gamma$. Then

$$\begin{aligned}
\int_{H/Z_H} \langle \pi(\gamma h)v, \tilde{\lambda}\rangle\ dh &= \int_{H/Z_H} \langle \pi(h)v, \tilde{v}_\gamma\rangle\ dh \\
&= \langle v, \tilde{v}_\gamma\rangle_\theta \\
&= d_\theta^{-1}(\pi)\ \langle v, \tilde{\lambda}\rangle\ \langle \lambda, \tilde{v}_\gamma\rangle \\
&= d_\theta^{-1}(\pi)\ \langle v, \tilde{\lambda}\rangle\ \Theta(\gamma).
\end{aligned}$$

Therefore, we obtain the desired formula. □

References

[BZ] I. N. Bernstein and A. V. Zelevinskii, Representations of the group GL(n, F) where F is a non-archimedean local field, Russian Math. Surveys 31:3 (1976), 1–68.

[H] J. Hakim, Supercuspidal Gelfand pairs, J. Number Theory **100** (2003), 251–269.

[HMa1] J. Hakim and Z. Mao, Cuspidal representations associated to $(GL(n), O(n))$ over finite fields and p-adic fields, J. Algebra **213** (1999), 129–143.

[HMa2] J. Hakim and Z. Mao, Supercuspidal representations of $GL(n)$ distinguished by a unitary subgroup, Pacific J. Math **185** (1998), no. 1, 149–162.

[HMu] J. Hakim and F. Murnaghan, Distinguished tame supercuspidal representations, preprint.

[HC] Harish-Chandra, A submersion principle and its applications, Papers dedicated to the memory of V. K. Patodi, Indian Academy of Sciences, Bangalore, and the Tata Institute of Fundamental Research, Bombay, 1980, pp. 95–102.

[KT] S.-I. Kato and K. Takano, Subrepresentation theorem for p-adic symmetric spaces, preprint.

[M1] F. Murnaghan, Representations of reductive p-adic groups, Fall 1997, unpublished course notes.

[M2] F. Murnaghan, Spherical characters: the supercuspidal case, Con. Math., to appear.

[RR] C. Rader and S. Rallis, Spherical characters on $\mathfrak{p}$-adic symmetric spaces, Amer. J. Math. 118 (1996), 91–178.

[RS] C. Rader and A. Silberger, Some consequences of Harish-Chandra's submersion principle, Proc. Amer. Math. Soc. 118 (1993), no. 4, 1271–1279.

J. Hakim
Department of Mathematics and Statistics,
American University,
4400 Massachusetts Ave NW,
Washington DC 20016, USA
Email: jhakim@american.edu

ALBANIAN JOURNAL
OF MATHEMATICS
Volume 1, Number 4, Pages 223–233
ISSN 1930-1235: (2007)

TORIC FIBRATIONS AND MIRROR SYMMETRY

ARTUR ELEZI

Abstract. The relation between the quantum $\mathcal{D}$-modules of a smooth variety X and a toric bundle is studied here. We describe the relation completely when X is a semi-ample complete intersection in a toric variety. In this case, we obtain all the relations in the small quantum cohomology ring of the bundle.

1. Introduction and Goals

For a smooth, projective variety Y we denote by $Y_{k,\beta}$ the moduli stack of rational stable maps of class $\beta \in H_2(Y,\mathbb{Z})$ with k-markings (Fulton et al [8]) and $[Y_{k,\beta}]$ its virtual fundamental class (Behrend et al [3], Li et al [13]). Genus zero Gromov-Witten invariants are defined as appropriate integrals over $[Y_{k,\beta}]$. We let $e : Y_{1,\beta} \to Y$ be the evaluation map, ψ - the first chern class of the cotangent line bundle on $Y_{1,\beta}$ and $\mathrm{ft} : Y_{1,\beta} \to Y_{0,\beta}$ - the forgetful morphism.
The formal completion of an arbitrary ring $\mathcal{R}$ along the semigroup MY of the rational curves of Y is defined to be

$$\mathcal{R}[[q^\beta]] := \{ \sum_{\beta \in \mathrm{MY}} a_\beta q^\beta, \quad a_\beta \in \mathcal{A}, \quad \beta - \text{effective}\}. \tag{1}$$

where $\beta \in H_2(Y,\mathbb{Z})$ is *effective* if it is a positive linear combination of rational curves. For each β, the set of α such that α and $\beta - \alpha$ are both effective is finite, hence $\mathcal{R}[[q^\beta]]$ behaves like a power series. Alternatively, we may define

$$q^\beta := q_1^{d_1} \cdot ... \cdot q_k^{d_k} = \exp(t_1 d_1 + ... + t_k d_k)$$

where $\{d_1, d_2, ..., d_k\}$ are the coordinates of β relative to the dual of a nef basis $\{p_1, ..., p_k\}$ of $H^2(Y,\mathbb{Q})$..
Let $*$ denote the small quantum product of Y. The small quantum cohomology ring

$$(QH_s^*Y, *)$$

is a deformation of the cohomology ring $(H^*(Y,\mathbb{Q}[q^\beta]), \cup)$. Its structural constants are three point Gromov-Witten invariants of genus zero. Let $\hbar$ be a formal variable and

$$J_\beta(Y) := e_* \left(\frac{[Y_{1,\beta}]}{\hbar(\hbar - \psi)} \right) = \sum_{k=0}^{\infty} \frac{1}{\hbar^{2+k}} e_*(\psi^k \cap [Y_{1,\beta}]).$$

2000 *Mathematics Subject Classification.* Primary 14N35, 53D45. Secondary: 14F05, 14J45, 14M25.

Key words and phrases. Gromov-Witten Theory, Quantum Cohomology, $\mathcal{D}$-module Structure, Mirror Theorems, Toric Bundle.

The sum is finite for dimension reasons. For $t = (t_0, t_1, ..., t_k)$, let

$$tp := t_0 + \sum_{i=1}^{k} t_i p_i.$$

The $\mathcal{D}$-module for the quantum differential equation of Y

$$1 \leq i \leq k,\ \hbar\partial/\partial t_i = p_i *,$$

is generated by (Givental [10])

$$J(Y) = \exp\left(\frac{tp}{\hbar}\right) \sum_{\beta \in H_2(Y,\mathbb{Z})} q^\beta J_\beta(Y)$$

where we use the convention $J_0 = 1$. The generator $J(Y)$ encodes *all* of the genus zero, one marking Gromov-Witten invariants and gravitational descendants of Y. The generator $J(Y)$ is an element of the completion $H^*(Y, \mathbb{Q})[t][[q^\beta]]$ that may be used to produce relations in QH_s^*Y in the following way: let

$$\mathcal{P}(\hbar, \hbar\partial/\partial t_i, q_i)$$

be a polynomial differential operator where q_i and $\hbar$ act via multiplication and $q_i = e^{t_i}$ are on the left of derivatives. If

$$\mathcal{P}(\hbar, \hbar\partial/\partial t_i, q_i)J(Y) = 0$$

then

$$\mathcal{P}(0, p_i, q_i) = 0$$

is a relation in the small quantum cohomology ring QH_s^*Y.
If Y is a complete intersection in a toric variety, $J(Y)$ is related to an explicit hypergeometric series $I(Y)$ via a change of variables (Givental [8], Lian et al [12],[13]). Furthermore, if Y is Fano then the change of variables is trivial, i.e.

$$J(Y) = I(Y).$$

Since $I(Y)$ is known explicitly, this yields two immediate benefits.

(1) The one point Gromov-Witten invariants and gravitational descendants of Y are determined completely.
(2) Differential operators that annihilate $I(Y)$ are easy to find, hence producing relations in the small quantum cohomology ring of Y.

In this paper we seek to relativize these results for Fano toric bundles, hence extending the results of the papers Elezi [6],[7]

2. Toric Bundles and Mirror Theorems

Toric varieties and bundles. We follow the approach and the terminology of Oda [15]. Let $\mathbb{M} \simeq \mathbb{Z}^m$ be a free abelian group of rank m, $\mathbb{N} = \mathrm{Hom}(\mathbb{M}, \mathbb{Z})$ its dual, and $<,>: \mathbb{M} \times \mathbb{N} \mapsto \mathbb{Z}$ the pairing between them. Let Y be an m-dimensional smooth, toric variety determined by a fan $\Sigma \subset \mathbb{N} \otimes \mathbb{R}$. Denote by

$$\Sigma(1) = \{\rho_1, ..., \rho_m, \rho_{m+1}, ..., \rho_{r=m+k}\}$$

the one dimensional cones of Σ and $D_1, ..., D_r$ the corresponding toric divisors. Let $v_i = (v_{i1}, ..., v_{im})$ be the first lattice point along the ray ρ_i. Let

$$\{a_1, a_2, ..., a_k\}$$

with $a_j := (a_{1j}, a_{2j}, ..., a_{mj}, a_{m+1j}, ..., a_{rj})$ be a basis of the lattice of relations Λ between $v_1, ..., v_r$. There is a short exact sequence

$$0 \to \Lambda \to \mathbb{Z}^{\Sigma(1)} \xrightarrow{h} \mathbb{N} \to 0, \tag{2}$$

where $h(c_1, c_2, ...c_r) = c_1v_1 + ... + c_rv_r$. The lattice Λ may be identified with $\mathrm{Hom}(A_{m-1}(Y), \mathbb{Z}) \simeq H_2(Y, \mathbb{Z})$. Under this isomorphism, a_{ij} is the intersection of a_j, when interpreted as a two dimensional cycle, with the toric divisor D_i. We choose a_j so that $\{a_1, ..., a_k\}$ is a generating set for the Mori cone of classes of effective curves. Then $a_{i1}, ..., a_{ik}$ are the coordinates of D_i with respect to the nef basis $\{p_1, ..., p_k\}$ dual to $\{a_1, ..., a_k\}$.
Assume that $\rho_1, ..., \rho_m$ generate a maximal dimensional cone in Σ. Since Y is smooth, $\{v_1, v_2, ..., v_m\}$ forms a $\mathbb{Z}$-basis of $\mathbb{N}$ and the absolute value of the matrix

$$(a_{ij});\ i = m+1, ..., r;\ j = 1, 2, ..., k$$

is 1.
The cohomology ring $H^*(Y, \mathbb{Z})$ is generated by the divisors $D_1, ..., D_r$ subject to the following two types of relations:
Type One: Whenever $\{\rho_{j_1}, ..., \rho_{j_s}\}$ do not generate a cone in Σ, the intersection

$$D_{j_1} \cdot ... \cdot D_{j_s} = 0. \tag{3}$$

Type Two: For each $1 \le i \le m$,

$$D_i = \sum_{j=1}^{k} a_{ij} p_j \tag{4}$$

From the short exact sequence (2) we obtain

$$0 \to \mathbb{T}^k \xrightarrow{\alpha} \mathbb{T}^r \xrightarrow{\beta} \mathbb{T}^m \to 0, \tag{5}$$

where the maps are defined as follows:

$$\alpha(t_1, t_2, ...t_k) = (\prod_{i=1}^{k} t_i^{a_{1i}}, ..., \prod_{i=1}^{k} t_i^{a_{ri}}),\ \beta(t_1, ...t_r) = (\prod_{i=1}^{r} t_i^{v_{i1}}, ..., \prod_{i=1}^{r} t_i^{v_{im}}).$$

Let $Z(\Sigma) \subset \mathbb{C}^r$ be the variety whose ideal is generated by the products of those variables which do *not* generate a cone in Σ. The toric variety Y is the geometric quotient (Cox [5])

$$\mathbb{C}^r - Z(\Sigma)//\mathbb{T}^k$$

where the torus acts as follows

$$t \cdot x = \left(\prod_{i=1}^{k} t_i^{a_{1i}} x_1, ..., \prod_{i=1}^{k} t_i^{a_{ri}} x_r \right). \tag{6}$$

The short exact sequence (5) yields an action of the quotient $\mathbb{T} := \mathbb{T}^m$ on Y.
The first chern class of the tangent bundle to Y is equal to

$$\sum_{i=1}^{r} D_i = \sum_{i=1}^{k} n_i p_i.$$

The toric variety Y is Fano iff $n_i > 0$ for all i.

We relativize the previous construction as follows. Consider the principal $\mathbb{T}$-bundle

$$\mathbb{E} := \oplus_{i=1}^{m}(L_i - \{0\}) \to X,$$

where L_i are line bundles over a smooth, projective variety X. Let $\mathbb{T}$ act fibrewisely on $\mathbb{E}$ and the diagonally on the first m-homogeneous coordinates of Y. The quotient space

$$Y(\mathbb{E}) := \mathbb{E} \times_{\mathbb{T}} Y$$

is a toric bundles over X with fiber isomorphic to Y. The bundle $Y(\mathbb{E})$ inherits a $\mathbb{T}$-action.
There is a projection map $\pi : Y(\mathbb{E}) \to Y$. The maximal cone generated by $\{\rho_1, \rho_2, ..., \rho_m\}$ determines a $\mathbb{T}$ fixed point q in Y whose homogeneous coordinates are $(0, 0, ..., 0, 1, 1, ...1)$. In the relativized setting, the $\mathbb{T}$-equivariant inclusion

$$q \hookrightarrow Y$$

yields a map

$$q(\mathbb{E}) \simeq X \overset{s}{\hookrightarrow} Y(\mathbb{E})$$

which is a section of π. This is also a fixed point component for the action of $\mathbb{T}$ on $Y(\mathbb{E})$. The other $\mathbb{T}$-fixed points of Y yield sections of π and these are all the fixed point components.
Toric divisors lift to divisors in $Y(\mathbb{E})$; these liftings will be denoted by the same letter in this paper. It was shown in Sankaran and Uma [17] that the two types of relations (3) and (4) lift in a natural way in $H^*(Y(\mathbb{E}), \mathbb{Z})$; namely

$$D_{j_1} \cdot ... \cdot D_{j_s} = 0$$

whenever $\{\rho_{j_1}, ..., \rho_{j_s}\}$ do not generate a cone in Σ, and

$$D_i = \sum_{j=1}^{k} a_{ij} p_j + c_1(L_i)$$

for each $1 \leq i \leq m$, where as in the case of $H^*(Y, \mathbb{Z})$ the divisors

$$p_1, ..., p_k$$

generate freely $H^*(Y(\mathbb{E}), \mathbb{Z})$. In fact, there is a simple relation between the $\mathbb{T}$-equivariant cohomology of Y and the cohomology of $Y(\mathbb{E})$ which will be used throughout this paper. Recall, that the rational cohomology of the classifying space $B\mathbb{T}$ is $\mathbb{Q}[\lambda_1, ...\lambda_m]$ where λ_i is the first chern class of the equivariant line bundle corresponding to the character

$$\nu_i : \mathbb{T} \to \mathbb{C}^* \ \nu_i(t_1, ..., t_m) = t_i.$$

A relation in the equivariant cohomology ring of Y becomes a relation in $H^*(Y(\mathbb{E}))$ after substituting $c_1(L_i)$ for λ_i.

We may assume that $L_i = \mathcal{O}_X, i > m$ without loss of generality. This is due to the fact that $\rho_1, ..., \rho_m$ generate a maximal cone in Σ.

The quantum $\mathcal{D}$-module structure of a toric bundle. The generator J of a quantum $\mathcal{D}$-structure is weighted by the lattice points of the Mori cone. Hence we first study the relation between the Mori cones of Y and $Y(\mathbb{E})$.

Lemma 1. *If L_i^* are generated by global sections, then the liftings of the nef divisors $p_1, ..., p_k$ in $Y(\mathbb{E})$ are also nef. Furthermore, the Mori cone of $Y(\mathbb{E})$ is a direct sum of the Mori cone of X, embedded via the section s, and the Mori cone of the fiber Y.*

Proof. In toric varieties, every nef divisor p is generated by global sections (Oda [14]). Let $x_1, x_2, ..., x_r$ be homogeneous coordinates in Y. The vector space of global sections $H^0(\mathcal{O}(p))$ has a monomial basis

$$\prod_{i=1}^{r} x_i^{m_i}.$$

Let $\{\phi_{ij}\}$ be a collection of generating sections for the line bundles L_i^*. The "monomials"

$$\prod_{i=1}^{r} (x_i \phi_{ij})^{m_i}$$

are generating sections the line bundle

$$\prod_{i=1}^{r} (\mathcal{O}(D_i) \otimes (L_i^*))^{m_i}$$

which is isomorphic to $\mathcal{O}(p)$ in $Y(\mathbb{E})$. Thus p lifts to a nef divisor in $Y(\mathbb{E})$.
This shows that the addition of $p_1, ..., p_k$ to a nef basis $\{p_{k+1}, ..., p_l\}$ of X yields a nef basis

$$\{p_1, ..., p_l\}$$

of $Y(\mathbb{E})$. Now for a curve $C \subset Y(\mathbb{E})$ we have

$$\pi_* \left([C] - s_*(\pi_*([C])) \right) = 0.$$

Notice that the restrictions of the divisors $p_1, p_2, ..., p_k$ in the section $q(\mathbb{E})$ are all zero since they may be written as $\mathbb{Z}$-linear combinations of $D_{m+1}, ..., D_{m+k}$. Hence $\forall i = 1, 2, ..., k,\ p_i \cdot ([C] - s_*(\pi_*([C]))) \geq 0$ and we have a unique decomposition

$$[C] = s_*(\pi_*([C])) + [C'],$$

where $[C']$ and $\pi_*([C])$ are curve classes respectively in the fiber of π and X. □

We introduce a "mixed" $I(Y(\mathbb{E}))$ that admits contributions from both $J(X)$ and an $\mathbb{E}$-twisted $J(Y)$. Let (ν, d) denote a curve class in the Mori cone of $Y(\mathbb{E})$, with ν a curve class in the fiber of π and d a curve class in X.
Define

$$I(Y(\mathbb{E})) := \exp\left(\frac{tp}{\hbar}\right) \sum_{(d,\nu)} {q_1}^{\nu} {q_2}^{d} \prod_{i=1}^{m} \frac{\prod_{m=0}^{\infty} (D_i + m\hbar)}{\prod_{m=0}^{D_i(\nu,d)} (D_i + m\hbar)} \pi^*(J_d(X)).$$

If X is a point then $Y(\mathbb{E}) = Y$. Furthermore, as mentioned in the introduction $J(Y) = I(Y)$ if Y is a Fano toric variety. In this paper we show that the same holds for the relativized $Y(\mathbb{E})$.

Proposition 1. *If X is a semi-ample complete intersection in a toric variety, and both Y and $Y(\mathbb{E})$ are Fano, then $J(Y(\mathbb{E})) = I(Y(\mathbb{E}))$.*

Proposition 1 will follow as a corollary of another statement which we now formulate and prove.
Let Z be a toric variety, $\tilde{L}_i,\ i=0,1,...,n$ toric line bundles over Z and $\tilde{\mathbb{E}}=\oplus_{i=0}^{n}\tilde{L}_i$. The bundle

$$\pi : Y(\tilde{\mathbb{E}}) \to Z$$

is also a toric variety (Oda [15]). The edges of the fan for $Y(\tilde{\mathbb{E}})$ corresponds to the liftings $B_1,...,B_r$ to $Y(\mathbb{E})$ of the toric base divisors $b_1,...,b_r$ and the divisors D_i from Y.
Let $\mathcal{L}_a : a=1,2,...,l$ be globally generated line bundles over Z and X the zero locus of a generic section s of

$$V=\oplus_{a=1}^{l}\mathcal{L}_a.$$

Such an X will be called *a semi-ample complete intersection*. Denote by L_i and $\mathbb{E}$ the restrictions of $\tilde{L}_i$ and $\tilde{\mathbb{E}}$ to X. The total space of $Y(\mathbb{E})$ is easily seen to be the zero locus of the section $\pi^*(s)$ of the pull back bundle $\pi^*(V)$.

Assume that the line bundles $\tilde{L}_i^*$ are globally generated and $-K_Z-\sum_{a=1}^{l}c_1(\mathcal{L}_a)+\sum_{i=0}^{n}c_1(\tilde{L}_i)$ is ample. (This will ensure that the conditions of Proposition 1 for the bundle $Y(\mathbb{E})$ over X are satisfied.)

Let V_d be the bundle on $Z_{1,d}$ whose fiber over the moduli point (C,x_1,f) is $\oplus_a H^0(f^*(\mathcal{L}_a))$. Denote by s_V its canonical section induced by s, i.e.

$$s_V((C,x_1,f))=f^*(s).$$

The stack theoretic zero section of s_V is the disjoint union

$$Z(s_V)=\coprod_{i_*(\beta)=d} X_{1,\beta}. \tag{7}$$

The map $i_* : H_2X \to H_2Z$ is not injective in general, hence the zero locus $Z(s_V)$ may have more then one connected component. An example is the quadric surface in $\mathbb{P}^3$. The sum of the virtual fundamental classes $[X_{1,\beta}]$ is the refined top Chern class of V_d with respect to s_V.
Let $\tilde{V}_{\nu,d}$ and $\tilde{s}_V$ be the pull backs of V_d and s_V via the stack morphism

$$Y(\tilde{\mathbb{E}})_{1,(\nu,d)} \to Z_{1,d}.$$

The zero section of $\tilde{s}_V$ is the disjoint union

$$z(\tilde{s}_V)=\coprod_{i_*(\beta)=d} Y(\mathbb{E})_{1,(\nu,\beta)}.$$

It follows that

$$\sum_{i_*(\beta)=d}[Y(\mathbb{E})_{1,(\nu,\beta)}]=c_{\text{top}}(\tilde{V}_{\nu,d})\cap[Y(\tilde{\mathbb{E}})_{1,(\nu,d)}].$$

Recall that the nef basis $\{p_1,p_2,...,p_k,p_{k+1},...p_l\}$ of $Y(\mathbb{E})$ is obtained by completing a nef basis $\{p_{k+1},...,p_l\}$ of X. We will use tp to denote both $\sum_{i=1}^{l}t_ip_i$ and $\sum_{i=k+1}^{l}t_ip_i$. The difference will be clear from the context.
Consider the following generating functions

$$J^V(Y(\tilde{\mathbb{E}}))=\exp\left(\frac{tp}{\hbar}\right)\sum_{(\nu,d)}{q_1}^{\nu}{q_2}^{d}e_*\left(\frac{c_{\text{top}}(\tilde{V}_{\nu,d})\cap[Y(\tilde{\mathbb{E}})_{1,(\nu,d)}]}{\hbar(\hbar-c)}\right)$$

and

$$\tilde{I}^V(Y(\tilde{\mathbb{E}})) = \exp\left(\frac{tp}{\hbar}\right) \sum_{(\nu,d)} {q_1}^{\nu} {q_2}^{d}\ \Omega_{\nu,d}\ \pi^* e_* \left(\frac{c_{\text{top}}(V_d) \cap [Z_{1,d}]}{\hbar(\hbar - c)}\right),$$

where

$$\Omega_{\nu,d} = \prod_{i=1}^{m} \frac{\prod_{m=0}^{\infty}(D_i + m\hbar)}{\prod_{m=0}^{D_i(\nu,d)}(D_i + m\hbar)}.$$

Proposition 2. *If* $-K_Y - \sum_{a=1}^{l} c_1(\mathcal{L}_a) - \sum_{i=0}^{n} c_1(\tilde{L}_i)$ *is ample then*

$$J^V((\tilde{\mathbb{E}})) = \tilde{I}^V(Y(\tilde{\mathbb{E}}))$$

Proof. Let

$$I_d^V(Z) = \prod_a \frac{\prod_{m=-\infty}^{\mathcal{L}_a(d)}(\mathcal{L}_a + m\hbar)}{\prod_{m=-\infty}^{0}(\mathcal{L}_a + m\hbar)} \prod_i \frac{\prod_{m=-\infty}^{0}(B_i + m\hbar)}{\prod_{m=-\infty}^{B_i(d)}(B_i + m\hbar)}.$$

From Givental [9], Lian et al [12], Lian et al [13] we know that $J^V(Y(\tilde{E}))$ is related via a mirror transformation to

$$I^V(Y(\tilde{\mathbb{E}})) = \exp\left(\frac{tp}{\hbar}\right) \cdot \sum {q_1}^{\nu} {q_2}^{d} \Omega_{\nu,d} I_d^V(Z).$$

Likewise

$$J^V(Z) = \exp\left(\frac{tp}{\hbar}\right) \sum {q_2}^{d} e_* \left(\frac{c_{\text{top}}(V_d) \cap [Z_{1,d}]}{\hbar(\hbar - c)}\right)$$

is related to

$$I^V(Z) = \exp\left(\frac{tp}{\hbar}\right) \sum {q_2}^{d} I_d^V(Z).$$

Since $-K_{Y(\tilde{E})} - \sum_a c_1(\mathcal{L}_a)$ and $-K_Z - \sum_a c_1(\mathcal{L}_a)$ are ample, the mirror transformations are particularly simple. Indeed, both series can be written as power series of $\hbar^{-1}$ as follows:

$$I^V(Y(\tilde{E})) = 1 + \frac{P_1(q_1, q_2)}{\hbar} + o(\hbar^{-1}),\ I^V(Z) = 1 + \frac{P_2(q_2)}{\hbar} + o(\hbar^{-1}),$$

where $P_1(q_1, q_2), P_2(q_2)$ are both polynomials supported respectively in

$$\Lambda_1 := \{(\nu, d) \mid (-K_{Y(\tilde{E})} - \sum c_1(\mathcal{L}_a)) = 1;\ D_j \geq 0,\ \forall j;\ B_i \geq 0,\ \forall i\},$$

and

$$\Lambda_2 := \{d \mid (-K_Z - \sum c_1(\mathcal{L}_a)) = 1; B_i \geq 0\ \forall i\}.$$

Then

$$J^V(Y(\tilde{E})) = \exp\left(\frac{-P_1(q_1, q_2)}{\hbar}\right) I^V(Y(\tilde{E}))$$

and

$$J^V(Z) = \exp\left(\frac{-P_2(q_2)}{\hbar}\right) I^V(Z).$$

Simple algebraic manipulations show that

- $c_1(\tilde{L}_j) \cdot d = 0, \forall d \in \Lambda_2, \forall j = 1, 2, ..., n$
- $\Lambda_1 = \{(0, d) \mid d \in \Lambda_2\}$.

It follows that $\Omega_{0,d} = 1, \forall d \in \Lambda_2$ hence $P_1(q_1, q_2) = P_2(q_2)$. Notice also that if we expand

$$\exp\left(\frac{-P_2(q_2)}{\hbar}\right) = \sum_{\alpha} c_\alpha q_2^\alpha$$

then

$$c_1(\tilde{L}_j) \cdot \alpha = 0, \forall j = 1, 2, ..., n.$$

Hence for each $(\nu, d) \in M\mathbb{P}(\tilde{V})$ we have $\Omega_{\nu,d} = \Omega_{\nu,d+\alpha}$. Now the proposition follows easily. □

Proof. of Proposition 1. We know return to the proof of Proposition 1. Recall that the map

$$i_* : H_2(X) \to H_2(Z) \tag{8}$$

is not necessarily injective in general. If it is, then

$$[X_{1,\beta}] = c_{\text{top}}(V_{i_*(\beta)}) \cap [Y_{1,i_*(\beta)}]$$

and

$$[Y(\mathbb{E})_{1,(\nu,\beta)}] = c_{\text{top}}(\tilde{V}_{\nu,i_*(\beta)}) \cap [Y(\tilde{\mathbb{E}})_{1,(\nu,i_*(\beta))}].$$

In this case one can easily show that

$$i_*(J_{\nu,\beta}(Y(\mathbb{E}))) = J^V_{\nu,i_*(\beta)}(Y(\tilde{\mathbb{E}}))$$

and

$$i_*(I_{\nu,\beta}(Y(\mathbb{E}))) = \tilde{I}^V_{\nu,i_*(\beta)}(Y(\tilde{\mathbb{E}})).$$

Proposition 2 shows that Proposition 1 holds for complete intersection in toric varieties for which the map (8) is injective.

□

3. Lifting the Quantum Cohomology Structure

In this section we use Proposition 1 to study small quantum cohomology ring of $Y(\mathbb{E})$. As explained in the introduction, some of the relations in the small quantum cohomology ring come from differential operators.

Proposition 3. *Whenever Proposition 1 holds, quantum differential operators of X may be lifted in $Y(\mathbb{E})$, while the quantum differential operators of the fiber Y may be extended to $Y(\mathbb{E})$. Both types of operators produce relations in the quantum cohomology $QH^*_s Y(\mathbb{E})$.*

Proof. Recall that $D_i = \sum a_{ij} p_j$. Let

$$c_1(L_i) = \sum_{j=k+1}^{l} c_{ij} p_j, \ i = 0, 1, ..., n.$$

Recall that the nef basis $\{p_1, p_2, ..., p_k, p_{k+1}, ...p_l\}$ of $H^2(Y(\mathbb{E}), \mathbb{Z})$ is obtained by completing a nef basis $\{p_{k+1}, ..., p_l\}$ of X. Let

$$\mathcal{P}(\hbar, \hbar\partial/\partial t_{k+1}, ..., \hbar\partial/\partial t_l, q_2) = \sum_{\alpha \in \Lambda} q_2^\alpha \mathcal{P}_\alpha$$

be a polynomial differential operator with Λ a finite subset of the Mori cone of X. Suppose that

$$0 = \mathcal{P}J(X) = \sum_{\alpha\in\Lambda} q_2^\alpha \sum_\beta \mathcal{P}_\alpha \left(\exp(\frac{pt}{\hbar})q_2^\beta\right) J_\beta(X)$$

$$= \sum_{\alpha\in\Lambda} q_2^\alpha \sum_\beta c_{\alpha,\beta}\exp(\frac{pt}{\hbar})q_2^\beta J_\beta(X) = \exp(\frac{pt}{\hbar}) \sum_{\alpha\in\Lambda,\beta} q_2^{\alpha+\beta} c_{\alpha,\beta} J_\beta(X).$$

Let

$$\delta_\alpha = \prod_{i=1}^n \prod_{r_i=0}^{-L_i\cdot\alpha-1} (\sum_{j=1}^k a_{ij}\hbar\frac{\partial}{\partial t_j} + \sum_{j=k+1}^l c_{ij}\hbar\frac{\partial}{\partial t_j} - r_i\hbar),\ \tilde{\mathcal{P}} = \sum_{\alpha\in\Lambda} q_2^\alpha \delta_\alpha \mathcal{P}_\alpha,$$

with the convention that if

$$L_i(\alpha) = 0,$$

the factors of δ_α corresponding to L_i are missing. Notice that

$$L_{n+1}(\alpha) = ... = L_m(\alpha) = 0$$

since we have chosen L_i to be trivial for $i > n$. We compute

$$\tilde{\mathcal{P}}J(Y(\mathbb{E})) = \sum_{\alpha\in\Lambda} q_2^\alpha \delta_\alpha \sum_{\nu,\beta} \mathcal{P}_\alpha \left(q_2^\beta \exp(\frac{pt}{\hbar})\right) q_1^\nu \Omega_{\nu,\beta} J_\beta =$$

$$\sum_{\alpha\in\Lambda} q_2^\alpha \delta_\alpha \sum_{\nu,\beta} c_{\alpha,\beta}\exp(\frac{pt}{\hbar}) q_1^\nu q_2^\beta \Omega_{\nu,\beta} J_\beta.$$

One can easily show that

$$\delta_\alpha \left(\exp(\frac{pt}{\hbar}) q_1^\nu q_2^\beta \Omega_{\nu,\beta}\right) = \exp(\frac{pt}{\hbar}) q_1^\nu q_2^\beta \Omega_{\nu,\alpha+\beta}.$$

It follows that

$$\tilde{\mathcal{P}}J(Y(\mathbb{E})) = \exp(\frac{pt}{\hbar}) \sum_\nu q_1^\nu \sum_{\alpha\in\Lambda,\beta} c_{\alpha,\beta} q_2^{\alpha+\beta} \Omega_{\nu,\alpha+\beta} J_\beta(X) = 0.$$

Hence the relation $\mathcal{P}(0, p_{k+1}, ..., p_l, q_2) = 0$ in QH_s^*X lifts into the relation

$$\mathcal{P}(0, p_{k+1}, ..., p_l, q_2 \prod_{i=1}^n D_i) = 0$$

in $QH_s^*Y(\mathbb{E})$, where

$$\left(\prod_{i=1}^n D_i\right)^\alpha := \prod_{i=1}^n D_i^{-L_i(\alpha)}, \forall \alpha \in MX.$$

For a curve class ν in the fiber of π, consider the following differential operator

$$\Delta_\nu(\hbar\frac{\partial}{\partial t_1}, ..., \hbar\frac{\partial}{\partial t_l}, q_j) := \prod_{i:D_i(\nu)>0} \prod_{m=0}^{D_i(\nu)-1} (\sum_{j=1}^k a_{ij}\hbar\frac{\partial}{\partial t_j} - \sum_{j=k+1}^l c_{ij}\hbar\frac{\partial}{\partial t_j} + m\hbar)$$

$$-q^\nu \prod_{i:D_i(\nu)<0} \prod_{m=0}^{-D_i(\nu)-1} (\sum_{j=1}^k a_{ij}\hbar\frac{\partial}{\partial t_j} - \sum_{j=k+1}^l c_{ij}\hbar\frac{\partial}{\partial t_j} + m\hbar).$$

It is easy to show that it satisfies

$$\Delta_\nu J(Y(\mathbb{E})) = 0.$$

It follows that

$$\Delta_\nu(p_1, ..., p_l, q_j) = 0$$

in $QH_s^*Y(\mathbb{E})$, i.e.

$$\prod_{i=1}^{r} D_i^{D_i(\nu)} = q^\nu.$$

These are precisely the extensions to $Y(\mathbb{E})$ of the small quantum cohomolgy relations of the fiber Y.

□

Sometimes *all* the relations in QH_s^*X come from quantum differential operators, hence QH_s^*X pulls back to $QH_s^*Y(\mathbb{E})$. This is the case when X is a Fano toric variety. The results of this section yield a complete description of $QH^*Y(E)$ which generalizes previous results of Costa et al [4] and Qin et al [15] and Givental [9].

4. The General (Nontoric) Case

We believe that Proposition 1 holds for any X. On one end, the equality of the $d = 0$ terms in $J(Y(\mathbb{E})) = I(Y(\mathbb{E}))$ is easy to establish. Indeed, the relative Gromov-Witten theory of the Y-bundle over $B\mathbb{T}$ associated with the universal bundle $E\mathbb{T} \mapsto B\mathbb{T}$ is precisely the $\mathbb{T}$-equivariant GW theory of Y (Astashkevich and Sadov [1]). The latter pulls back under the classifying map $X \mapsto B\mathbb{T}$ to the relative GW theory of $Y(\mathbb{E})$ over X. It follows that the restriction of $J(Y(\mathbb{E}))$ to $\nu = 0$ is obtained by substituting $c_1(L_i)$ for λ_i in $J^{\mathbb{T}}(Y)$. Since Y is assumed to be Fano, the generator $J^{\mathbb{T}}(Y)$ is known (see for example [8]) and the substitution $c_1(L_i) \mapsto \lambda_i$ is easily seen to yield the desired equality. At the other end, the $\nu = 0$ equality follows as an application of the equivariant quantum Lefshetz principle for the action of a torus on the fibers of $Y(\mathbb{E})$. The fixed point component relevant for the equivariant and localization considerations ([12]) consists of the maps that land in the section $s(X)$. The top chern class of the virtual normal bundle for this component is that of the $\mathbb{H}^1$-bundle for $\oplus_{i=1}^m L_i$. Calculations are easy to carry out (see for example Elezi [7]).

References

[1] A. Astashkevich and V. Sadov, *Quantum Cohomology of Partial Flag manifolds*, Comm. Math. Phys. **170** (1995) no 3, 503-528.

[2] M. Atiyah and R. Bott, *The moment map and equivariant cohomology*, Topology **23**, 1-28.

[3] K. Behrend and B. Fantechi, *The intrinsic normal cone*, Invent. Math. 128(1): 45-88, 1997.

[4] L. Costa and R. M. Mir-Roig, *Quantum cohomology of projective bundles over* $\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_s}$, Internat. J. Math. 11 (2000), **6**, 761-797.

[5] D. Cox, *Homogeneous Coordinate Ring of a Toric Variety*, J. Algebraic Geom. **4** 1995, 17-50.

[6] A. Elezi, *A Mirror Conjecture for Projective Bundles*, Internat. Math. Res. Notices **55** 2005, 3445-3458.

[7] A. Elezi, *Mirror Symmetry and Quantum Cohomology for Projective Bundles*, Int. J. Pure Appl. Math. (2007) **36**, no. 1, 75-86.

[8] W. Fulton and R. Pandharipande, *Notes on stable maps and quantum cohomology*, in *Proceedings of symposia in pure mathematics: Algebraic geometry Santa Cruz 1995*, **62** Part 2, 45-96.

[9] A. Givental, *A mirror theorem for toric complete intersections*, in *Topological field theory, primitive forms and related topics (Kyoto, 1996)*, Progr. Math., **160**, Birkhăuser, 1998 141-175.
[10] A. Givental, *Equivariant Gromov-Witten invariants*, Int. Math. Res. Notices **13** (1996), 613-663.
[11] A. Givental, *Homological Geometry and Mirror Symmetry*, in *Proceedings of the ICM*, Zürich 1994, Birkhäuser, 1995, v.1, 472-480.
[12] B. Lian, K. Liu, and S.-T.Yau, *Mirror principle II*, Asian J. Math **3** (1999), no. 1, 109-146.
[13] B. Lian, K. Liu, and S.-T.Yau, *Mirror principle III*, Asian J. Math, **3** (1999), no. 4, 771-800.
[14] J. Li and G. Tian, *Virtual moduli cycles and Gromov-Witten invariants of algebraic varieties*, Jour. AMS 11 (1998), no 1, 119-174.
[15] T. Oda, *Convex Bodies and Algebraic Geometry*, Springer-Verlag, Berlin, 1998.
[16] Zh. Qin and Y. Ruan, *Quantum cohomology of projective bundles over* $\mathbf{P}^n$, Trans. Amer. Math. Soc. 350 (1998), no. 9, 3615–3638.
[17] P. Sankaran and V. Uma, *Cohomology of toric bundles*, Comm. Math. Helv. **78**, (2003), 540-554.

A. Elezi
Department of Mathematics and Statistics,
American University,
4400 Massachusetts Ave NW,
Washington DC 20016, USA
Email: aelezi@american.edu

ALBANIAN JOURNAL
OF MATHEMATICS
Volume 1, Number 4, Pages 235–244
ISSN 1930-1235: (2007)

A VANISHING RESULT FOR TORIC VARIETIES ASSOCIATED WITH ROOT SYSTEMS

QËNDRIM R. GASHI

Abstract. Consider a root system R and the corresponding toric variety V_R whose fan is the Weyl fan and whose lattice of characters is given by the root lattice for R. We prove the vanishing of the higher cohomology groups for certain line bundles on V_R by proving a purely combinatorial result for root systems. These results are related to a converse to Mazur's Inequality for (simply-connected) split reductive groups.

1. Introduction

The problem of p-adically estimating the number of points on an algebraic variety over a finite field of characteristic p is quite old. An answer to a special case of this problem is given by the classical theorem of C. Chevalley and E. Warning, which asserts that if $G(x_1, ..., x_n)$ is a polynomial of degree less than n with integral coefficients, then the number of roots of

$$G(x_1, ..., x_n) \equiv 0(mod\ p)$$

is divisible by p.

N. Katz conjectured a sharper p-adic estimate for the number of solutions to the above equation, which then B. Mazur proved in [12] (and later P. Berthelot and A. Ogus completed in [1]), making use of the then recently discovered crystalline cohomology. Mazur's result is now frequently referred to as Mazur's Inequality and it is most easily stated using the so-called Newton and Hodge vectors. We will recall this inequality in the setting of F-isocrystals.

An F-isocrystal is a pair (N, F), where N is a finite-dimensional vector space over the fraction field K of the ring of Witt vectors $W(\overline{\mathbb{F}}_p)$, equipped with a Frobenius-linear bijective endomorphism F of N.

Suppose now that our isocrystal (N, F) is n-dimensional. If M is a $W(\overline{\mathbb{F}}_p)$-lattice in N, then we can associate to it the Hodge vector $\mu(M) \in \mathbb{Z}^n$, which measures the relative position of the lattices M and FM. By Dieudonné-Manin theory, we can associate to N its Newton vector $\nu(N, F) \in \mathbb{Q}^n$, which classifies F-isocrystals of dimension n up to isomorphism.

If $\geq$ stands for the "usual dominance order", then Mazur's Inequality asserts that $\mu(M) \geq \nu(N, F)$. In other words, if $\mu(M) = (\mu_1, ..., \mu_n)$ and $\nu(M) = (\nu_1, ..., \nu_n)$, then $\mu_1 \geq \nu_1$, $\mu_1+\mu_2 \geq \nu_1+\nu_2$,..., $\mu_1+...+\mu_{n-1} \geq \nu_1+...+\nu_{n-1}$, and $\mu_1+...+\mu_n = \nu_1 + ... + \nu_n$.

R. Kottwitz and M. Rapoport in [10] proved a converse to this inequality. Namely, they proved that if we let (N, F) be an isocrystal of dimension n, and

let $\mu = (\mu_1, ..., \mu_n) \in \mathbb{Z}^n$, with $\mu_1 \geq \mu_2 \geq ... \geq \mu_n$, be such that $\mu \geq \nu(N, F)$, then there exists a $W(\overline{\mathbb{F}}_p)$-lattice M in N satisfying $\mu = \mu(M)$.

Both Mazur's Inequality and its converse can be regarded as statements for the group GL_n—the dominance order arises naturally in the context of the root system for GL_n. Kottwitz and Rapoport in [10] (see also [5] and [14]) formulated a group-theoretic version of the above converse to Mazur's Inequality, which they proved for GL_n and GSp_{2n}. They also reduced this problem to a combinatorial-type problem formulated purely in terms of root systems (see e.g. [8] for an explicit statement). Then C. Lucarreli in [11], proved that combinatorial statement for all (split connected) classical groups.

In [7], a new interpretation for the root-system combinatorial problem mentioned above was introduced; it was shown that it is equivalent to the vanishing of higher cohomology groups for certain line bundles on toric varieties associated with root systems. In loc.cit. a generalized version of this was proved for GL_n and the usual version for G_2, in particular giving an easy proof for the converse to Mazur's Inequality for GL_n and G_2 respectively. One of the surprising outcomes of these results, apart from the link of the converse to Mazur's Inequality to toric varieties, is that they improve the classical vanishing theorems for these toric varieties (compare, for example, results in [13]). They also give new ideas for new vanishing theorems for general toric varieties.

In the current paper, we take a new approach to proving the combinatorial problem mentioned above (see Theorem 2.1 below). Also, unlike in other works related to this problem that have appeared so far, all the root systems are considered at once and the method of the proof applies to each of them. We prove a vanishing result (see Theorem 2.4 below) for toric varieties associated with *any* root system. Also, we believe that these results provide a crucial step and a clear strategy for proving the converse of Mazur's Inequality for all (split, connected) groups.

Finally, we mention that these toric varieties have appeared in the De Concini-Procesi theory of group compactifications (see e.g. [2], pp.187–206) and are still actively used and studied in that field. They have also appeared on the work related to the local trace formula (see e.g. [9], §23) and the fundamental lemma (see e.g. [4], §5).

Acknowledgments: The author thanks his thesis adviser, Professor Robert Kottwitz, for comments on an earlier draft of this paper, and for continuous encouragement and support.

2. Set-up and Results

We follow the terminology from [3]. Let R be an irreducible, reduced root system and let $Q(R)$ stand for the root lattice for R. Denote by W_R the Weyl group for R. Let $x \in Q(R)$ and consider $O_x := \{wx : w \in W_R\}$, the Weyl orbit of x; write $\mathrm{Conv}(O_x)$ for the convex hull of O_x in $Q(R) \otimes_{\mathbb{Z}} \mathbb{R}$. Fix a root α in R and denote by $\alpha^\vee$ the corresponding coroot. Suppose that $z = y + \frac{1}{2}\alpha \in Q(R)$ is such that $y \in \mathrm{Conv}(O_x)$ and $\langle y, \alpha^\vee \rangle = 0$.

The main result of this paper, formulated combinatorially, is the following:

Theorem 2.1. *With x and z as above, we have that $z \in \mathrm{Conv}\,(O_x)$.*

We prove this theorem in the next section but in the rest of this section we briefly explain how it is related to a converse to Mazur's inequality and to toric varieties

associated with root systems (for more details on this relationship as well as for detailed references see e.g. [7], Introduction and Section 1 in particular).

Once a notation is introduced, it will be fixed for the rest of the paper. For a short time we will not be working over complex numbers.

Let F be a finite extension of $\mathbb{Q}_p$. Denote by $\mathcal{O}_F$ the ring of integers of F. Suppose G is a split, connected reductive group, B a Borel subgroup, and T a maximal torus in B, all defined over $\mathcal{O}_F$. Let $P = MN$ be a parabolic subgroup of G which contains B, where M is the unique Levi subgroup of P containing T.

We write X for the set of cocharacters $X_*(T)$. Then X_G and X_M will stand for the quotient of X by the coroot lattice for G and M, respectively. Also, we let $\varphi_G : X \to X_G$ and $\varphi_M : X \to X_M$ denote the respective natural projection maps.

Let $\mu \in X$ be G-dominant and let W be the Weyl group of T in G. The group W acts on X and so we consider $W\mu := \{w(\mu) : w \in W\}$ and the convex hull of $W\mu$ in $\mathfrak{a} := X \otimes_{\mathbb{Z}} \mathbb{R}$, which we denote by $\mathrm{Conv}\,(W\mu)$. Define

$$P_\mu = \{\nu \in X : (i)\, \varphi_G(\nu) = \varphi_G(\mu);\ \textit{and}\ (ii)\, \nu \in \mathrm{Conv}\,(W\mu)\}\,.$$

Let $\mathfrak{a}_M := X_M \otimes_{\mathbb{Z}} \mathbb{R}$ and write $pr_M : \mathfrak{a} \to \mathfrak{a}_M$ for the natural projection induced by φ_M. Note that X_M is a quotient of X, but we can consider $\mathfrak{a}_M$ as a subspace of $\mathfrak{a}$ (after tensoring with $\mathbb{R}$ any possible torsion is lost).

We now recall an important conjecture of Kottwitz and Rapoport. See e.g. [8], sections 4.3 and 4.4, where it is explained how a converse to Mazur's inequality follows from this conjecture.

Conjecture 2.2. *(Kottwitz-Rapoport) Let G be a split, connected, reductive group over F. Keeping the same notation as above, we have*

$$\varphi_M\,(P_\mu) = \{\nu_1 \in X_M : (i)\, \nu_1, \mu \text{ have the same image in } X_G;\\ (ii)\, \text{the image of } \nu_1 \text{ in } \mathfrak{a}_M \text{ lies in } pr_M\,(\mathrm{Conv}\, W\mu)\}\,.$$

It is worth mentioning that C. Lucarelli (see e.g. [11], Theorem 0.2) has proved that this conjecture holds for every split, connected reductive group over F where every irreducible component of its Dynkin diagram is of type A_n, B_n, C_n, or D_n. In [7], this has been further extended to include the group G_2 and a strengthened result has been proved in the case of the group GL_n (see theorems A and B in loc. cit.). We again refer the reader to the Introduction in [7] where references are given for earlier work on this conjecture by other mathematicians.

In this set-up, our main result can be reformulated as follows.

Theorem 2.3. *The above conjecture of Kottwitz and Rapoport is true for all split, connected, simply-connected semisimple groups in the case when the parabolic subgroup P is of semisimple rank 1.*

We would like to make a comment about some advantages and disadvantages of the proof of Theorem 2.1 (and therefore Theorem 2.3) . One of the advantages of the proof presented in this paper is that, unlike the previous proofs of (parts of) the Kottwitz-Rapoport Conjecture, all the root systems are considered at once and the method of the proof applies to each of them. Therefore, in particular, one gets a clearer picture, at least on the root-system level, of the entire problem of a converse to Mazur's inequality. Our approach has some serious disadvantages, however. The assumption that the group G is simply-connected is used in an

essential way (equivalently, we only consider the root lattice and not the weight lattice for our corresponding root system). Also, we only deal with the case when the parabolic subgroup P has semisimple rank 1. Admittedly, the latter is less of a "serious" problem and, in fact, the author believes that it can be overcome without introducing new strategies in the proof.

Let us now very briefly recall how this is all connected to vanishing results on certain toric varieties. The reader is encouraged to consult [7] (especially the Introduction and Section 1 therein) for more details.

From now on we assume that G is simply-connected. Then $\hat{G}$ is adjoint. Let $\hat{G}$ and $\hat{T}$ be the (Langlands) complex dual group for G and T, respectively. Then the corresponding toric variety, which we denote by V_G, is given by requiring that its fan be given by the Weyl fan in $X_*(\hat{T}) \otimes_{\mathbb{Z}} \mathbb{R}$ and its torus be $\hat{T}$. (We would like to remark that the toric variety V_G appears naturally in the theory of group compactifications—see e.g. [2], pp.187–206.) Note that it is now safe for the reader to assume that we are working over complex numbers.

Using the canonical surjection, we define a map

$$pr_M : X^*(\hat{T}) \twoheadrightarrow X^*(\hat{T})/R_{\hat{M}},$$

where $R_{\hat{M}}$ stands for the root lattice for $\hat{M}$ (note that the codomain of the last map is just $X^*(Z(\hat{M}))$).

For M as above (a Levi subgroup containing T), we will need a toric variety Y_M^G for the torus $Z(\hat{M})$ (see e.g. [9], §23.2). Note that $Z(\hat{M})$ is a subtorus of $\hat{T}$ and so $X_*(Z(\hat{M}))$ is a subgroup of $X_*(\hat{T})$. The collection of cones from the Weyl fan inside $X_*(\hat{T}) \otimes_{\mathbb{Z}} \mathbb{R}$ that lie in the subspace $X_*(Z(\hat{M})) \otimes_{\mathbb{Z}} \mathbb{R}$ gives a fan. This is the fan for the nonsingular, projective toric variety Y_M^G.

Let us now assume that our parabolic subgroup P is of semisimple rank 1. This implies that the root lattice $R_{\hat{M}}$ is just $\mathbb{Z}\alpha$ for a unique, up to a sign, root α of $\hat{G}$, and that the toric variety Y_M^G, which we now denote by D_α, is a non-torus-invariant divisor in V_G. The map pr_M will now be denoted by p_α. By tensoring with $\mathbb{R}$ we get a map from p_α, which we still denote by

$$p_\alpha : X^*(\hat{T}) \otimes_{\mathbb{Z}} \mathbb{R} \twoheadrightarrow (X^*(\hat{T})/\mathbb{Z}\alpha) \otimes_{\mathbb{Z}} \mathbb{R}.$$

Since tensoring with $\mathbb{R}$ will lose any possible torsion, we can identify the codomain of the last map with the coroot hyperplane

$$[\alpha^\vee = 0] := \{x \in X^*(\hat{T}) \otimes_{\mathbb{Z}} \mathbb{R} : \langle \alpha^\vee, x \rangle = 0\},$$

where $\langle , \rangle$ is the canonical pairing between cocharacters and characters, and $\alpha^\vee$ stands for the coroot of $\hat{G}$ corresponding to α. Also, the fan of D_α is contained in the root hyperplane

$$[\alpha = 0] := \{x \in X_*(\hat{T}) \otimes_{\mathbb{Z}} \mathbb{R} : \langle x, \alpha \rangle = 0\}.$$

Now let $\mathcal{L}$ be a $\hat{T}$-equivariant line bundle on V_G that is generated by its sections. Also, assume that $\mathcal{L}$ is invariant under the obvious action of W. Then we have a short-exact sequence of sheaves on V_G:

$$0 \longrightarrow \mathcal{J}_{D_\alpha} \otimes \mathcal{L} \longrightarrow \mathcal{L} \longrightarrow i_*(\mathcal{L}|_{D_\alpha}) \longrightarrow 0,$$

where $\mathcal{J}_{D_\alpha}$ is the ideal sheaf of D_α and i is the inclusion map $D_\alpha \hookrightarrow V_G$. Note that $H^i(V_G, \mathcal{L})$=0, for all $i > 0$, since $\mathcal{L}$ is generated by its sections and V_G is projective,

and also

$$H^i(V_G, i_*(\mathcal{L}|_{D_\alpha})) = H^i(D_\alpha, \mathcal{L}|_{D_\alpha}) = 0,$$

for all $i > 0$, since $\mathcal{L}|_{D_\alpha}$ is generated by its sections and D_α is a projective toric variety. Therefore the short-exact sequence gives rise to the long-exact sequence

$$... \longrightarrow H^0(V_G, \mathcal{L}) \xrightarrow{\varphi} H^0(V_G, i_*(\mathcal{L}|_{D_\alpha})) \longrightarrow H^1(V_G, \mathcal{J}_{D_\alpha} \otimes \mathcal{L}) \longrightarrow 0.$$

Using the usual combinatorial interpretation of the 0-th cohomology (see e.g. [6], pg. 66) we see that the map φ is induced by the map p_α. Also, clearly, the surjectivity of φ is equivalent to $H^1(V_G, \mathcal{J}_{D_\alpha} \otimes \mathcal{L}) = 0$. Then one finds (as is explained in Section 1 of [7]) that φ is surjective if and only if Theorem 2.3 is true. Thus, in this language, our result can be written as follows.

Theorem 2.4. *With notation as above, we have that*

$$H^i(V_G, \mathcal{J}_{D_\alpha} \otimes \mathcal{L}) = 0\,, \forall i > 0.$$

Before we end this section it is worth mentioning that vanishing results like the one in Theorem 2.4 are also of independent interest just from a toric-variety point of view. A very important vanishing result for toric varieties has been proved by Mustaţă (see e.g. [13], Theorem 2.1), and for the toric varieties associated with root systems, Theorem 2.4 gives the vanishing of higher cohomology groups for more line bundles on these varieties.

As we mentioned earlier, the next section is devoted to proving Theorem 2.1.

3. Proof of Theorem 2.1

We recall some of the notation from the previous section. We follow the terminology and notation from [3]. Let R be an irreducible, reduced root system and let $Q(R)$ stand for the root lattice for R. Suppose $\{\alpha_i : i \in I\}$ is the set of simple roots (for some choice of a chamber for R) and let s_{α_i} be the simple reflection corresponding to α_i, $i \in I$ (we will assume that $I := \{1, \ldots, n\}$ for some natural number n); then $\alpha_i^\vee$, $i \in I$, stands for the coroot corresponding to α_i. Moreover, let $\{\varpi_i : i \in I\}$ denote the set of fundamental coweights, where $\langle \varpi_i, \alpha_i \rangle = 1$, and $\langle , \rangle$ is the standard pairing between coweights and weights for R. We write W for the Weyl group of R (note that earlier we were writing W_R instead).

Let $x \in Q(R)$ and consider $O_x := \{wx : w \in W\}$, the Weyl orbit of x; write $\mathrm{Conv}(O_x)$ for the convex hull of O_x in $Q(R) \otimes_{\mathbb{Z}} \mathbb{R}$. Fix a root α in R and denote by $\alpha^\vee$ the corresponding coroot. Suppose that $z = y + \frac{1}{2}\alpha \in Q(R)$ is such that $y \in \mathrm{Conv}(O_x)$ and $\langle y, \alpha^\vee \rangle = 0$.

We would like to prove that $z \in \mathrm{Conv}(O_x)$. But, before we start the proof, we make a few reductions. First, by possibly choosing a different chamber of R, we can assume that α is a simple root, and we will therefore denote the latter by α_{i_0}, where i_0 is some element in I. Second, it is clear that we may take x to be dominant, since we are considering the convex hull of its orbit O_x under W, and for some $w \in W$ we must have that $wx \in O_x$ is dominant. Finally, we claim that it suffices to prove the theorem under the assumption that y is dominant; indeed, if y is not dominant, then we can find $w \in W$ such that wy is dominant, and we can apply the theorem to wy to find that $wz \in \mathrm{Conv}(O_x)$, and therefore $z \in \mathrm{Conv}(O_x)$.

Hence, to prove Theorem 2.1 it suffices to prove the following lemma.

Lemma 3.1. *Let $x \in Q(R)$ be dominant. Fix a simple root $\alpha_{i_0} \in S$ and suppose that $z = y + \frac{1}{2}\alpha_{i_0} \in Q(R)$ is such that*

(i) $\langle y, \alpha_{i_0}^\vee \rangle = 0$
(ii) *y is dominant*
(iii) $\langle y, \varpi_i \rangle \leq \langle x, \varpi_i \rangle$, $\forall i \in I$.

Then we have that $z \in \mathrm{Conv}\,(O_x)$.

(Since y is dominant, the condition (iii) in the lemma is equivalent to the assumption $y \in \mathrm{Conv}(O_x)$.)

The idea of the proof is as follows. First, notice that

$$\mathrm{Conv}(O_x) = \{u \in Q(R)_{\mathbb{R}} \,|\langle wu, \varpi_i \rangle \leq \langle x, \varpi_i \rangle,\ \forall w \in W,\ \forall i \in I\}.$$

In fact, we know that there exists a unique $w_0 \in W$ such that $w_0 z$ is dominant, and since x is assumed to be dominant and $\mathrm{Conv}(O_x)$ is invariant under the action of W, we get that

$$z \in \mathrm{Conv}(O_x) \Longleftrightarrow \langle w_0 z, \varpi_i \rangle \leq \langle x, \varpi_i \rangle,\ \forall i \in I. \tag{1}$$

Thus, the idea of the proof will be to find an element $w_0 \in W$ so that $w_0 z$ is dominant and then to prove that the inequalities on the right-hand side of (1) are satisfied.

Now we begin with the proof itsef. From the way z was defined, we see that

$$\langle z, \varpi_i \rangle = \langle y, \varpi_i \rangle,\ \forall i \in I \setminus \{i_0\},$$

and

$$\langle z, \varpi_{i_0} \rangle = \langle y, \varpi_{i_0} \rangle + \frac{1}{2}.$$

Then condition (iii) immediately gives

$$\langle z, \varpi_i \rangle \leq \langle x, \varpi_i \rangle,\ \forall i \in I \setminus \{i_0\}.$$

And, since $z \in Q(R)$, or more precisely since $\langle z, \varpi_{i_0} \rangle$ is an integer (this is where we are using the fact that we are working with the root lattice and not the weight lattice!), the relation (iii) also gives $\langle z, \varpi_{i_0} \rangle \leq \langle x, \varpi_{i_0} \rangle$. Thus

$$\langle z, \varpi_i \rangle \leq \langle x, \varpi_i \rangle,\ \forall i \in I. \tag{2}$$

In the case when z is dominant (or equivalently when $w_0 = 1$), the last set of inequalities implies that the inequalities in (1) are satisfied and so the assertion of our lemma is true.

But z need not be dominant. We have

$$\langle z, \alpha_i^\vee \rangle = \langle y, \alpha_i^\vee \rangle + \frac{1}{2}\langle \alpha_{i_0}, \alpha_i^\vee \rangle,\ \forall i \in I.$$

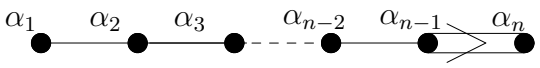
and (i) implies that $\langle z, \alpha_{i_0}^\vee \rangle = 1$. However, y being dominant, and having $\langle z, \alpha_i^\vee \rangle \in \mathbb{Z}$, we find that z is not dominant if and only if

$$\exists\, i_1 \in I \setminus \{i_0\} \,\text{s.t.}\, \langle \alpha_{i_0}, \alpha_{i_1}^\vee \rangle < -1 \,\text{and}\, \langle z, \alpha_{i_1}^\vee \rangle = -1. \tag{3}$$

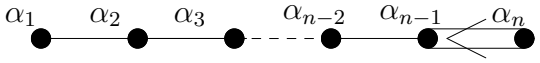
Note that if i_1 as above exists, then it is unique. Clearly, for $i \in I \setminus \{i_1\}$, $\langle z, \alpha_i^\vee \rangle \geq 0$.

Since $\langle \alpha_{i_0}, \alpha_{i_1}^\vee \rangle < -1$ can only happen for non-simply-laced root systems, we can conclude that z is always dominant for a simply-laced root system R and, by what we wrote above, the Lemma is true for such an R.

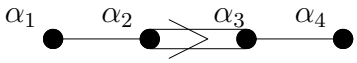
More specifically, if we use the notation as in the diagrams below, (3) holds only if:



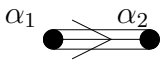
$R = B_n$ and $\alpha_{i_0} = \alpha_{n-1}$, $\alpha_{i_1} = \alpha_n$, and $\langle z, \alpha_{i_1}^\vee \rangle = -1$; or



$R = C_n$ and $\alpha_{i_0} = \alpha_n$, $\alpha_{i_1} = \alpha_{n-1}$, and $\langle z, \alpha_{i_1}^\vee \rangle = -1$; or



$R = F_4$ and $\alpha_{i_0} = \alpha_2$, $\alpha_{i_1} = \alpha_3$, and $\langle z, \alpha_{i_1}^\vee \rangle = -1$; or



$R = G_2$ and $\alpha_{i_0} = \alpha_1$, $\alpha_{i_1} = \alpha_2$, and $\langle z, \alpha_{i_1}^\vee \rangle = -1$.

Let us now assume that (3) holds. Our aim is to find $w_0 \in W$ so that $w_0 z$ is dominant. We first apply the simple reflection $s_{\alpha_{i_1}}$ to bring z "closer" to the dominant Weyl chamber. If we write $z_1 := s_{\alpha_{i_1}}(z)$, we see that $z_1 = z + \alpha_{i_1}$, so

$$\langle z_1, \alpha_i^\vee \rangle = \langle z, \alpha_i^\vee \rangle + \langle \alpha_{i_1}, \alpha_i^\vee \rangle, \, \forall i \in I,$$

and it is also easy to check that $\langle z_1, \alpha_i^\vee \rangle \geq 0$, for $i = i_0, i_1$.

We then find that z_1 is not dominant if and only if

$$\exists \, i_2 \in I \setminus \{i_0, i_1\} \text{ s.t. } \langle \alpha_{i_1}, \alpha_{i_2}^\vee \rangle = -1, \text{ and } \langle z, \alpha_{i_2}^\vee \rangle = 0. \tag{4}$$

Let us remark that if i_2 as above exists, then it must be unique. Clearly, for $i \in I \setminus \{i_2\}$, $\langle z_1, \alpha_i^\vee \rangle \geq 0$.

Note that if $R = G_2$, then z_1 is always dominant, because if (4) holds then α_{i_2} must be different from both α_{i_0} and α_{i_1}, but this cannot happen since G_2 has only two simple roots. So, for G_2, $w_0 = s_{\alpha_{i_1}}$.

For $R = B_n$, just as in the case of G_2, z_1 must be dominant, because there is no simple root α_{i_2}, distinct from α_{i_0}, such that $\langle \alpha_{i_1}, \alpha_{i_2}^\vee \rangle = -1$. So, for B_n, $w_0 = s_{\alpha_{i_1}}$.

If one uses the same notation as in the Dynkin diagrams above, then one can prove that the only possibility for z_1 to be non-dominant is if:

$R = C_n$ and $\alpha_{i_0} = \alpha_n$, $\alpha_{i_1} = \alpha_{n-1}$, $\alpha_{i_2} = \alpha_{n-2}$, $\langle z, \alpha_{i_1}^\vee \rangle = -1$, $\langle z, \alpha_{i_2}^\vee \rangle = 0$;
or
$R = F_4$ and $\alpha_{i_0} = \alpha_2$, $\alpha_{i_1} = \alpha_3$, $\alpha_{i_2} = \alpha_4$, $\langle z, \alpha_{i_1}^\vee \rangle = -1$, $\langle z, \alpha_{i_2}^\vee \rangle = 0$.

Now we assume that (3) and (4) hold. We write $z_2 := s_{\alpha_{i_2}}(z_1)$, so $z_2 = z + \alpha_{i_1} + \alpha_{i_2}$, and therefore

$$\langle z_2, \alpha_i^\vee \rangle = \langle z, \alpha_i^\vee \rangle + \langle \alpha_{i_1}, \alpha_i^\vee \rangle \; + \langle \alpha_{i_2}, \alpha_i^\vee \rangle, \, \forall i \in I.$$

It is easy to see that $\langle z_2, \alpha_i^\vee \rangle \geq 0$, for $i = i_0, i_1, i_2$. We then get that z_2 is not dominant if and only if

$$\exists \, i_3 \in I \setminus \{i_0, i_1, i_2\} \text{ s.t. } \langle \alpha_{i_2}, \alpha_{i_3}^\vee \rangle = -1, \text{ and } \langle z, \alpha_{i_3}^\vee \rangle = 0. \tag{5}$$

Note that if i_3 as above exists, then it is unique. Clearly, for $i \in I \setminus \{i_3\}$, $\langle z_2, \alpha_i^\vee \rangle \geq 0$.

Let us remark that this way we find that for $R = F_4$ we have $w_0 = s_{\alpha_{i_2}} \circ s_{\alpha_{i_1}}$, because there is no α_{i_3} that satisfies (5). So, it only remains to find a w_0 for $R = C_n$.

In any case, if z_2 is not dominant, we then apply $s_{\alpha_{i_3}}$ to it, to get $z_3 := s_{\alpha_{i_3}}(z_2) = z + \alpha_{i_1} + \alpha_{i_2} + \alpha_{i_3}$, and hence

$$\langle z_3, \alpha_i^\vee \rangle = \langle z, \alpha_i^\vee \rangle + \langle \alpha_{i_1}, \alpha_i^\vee \rangle + \langle \alpha_{i_2}, \alpha_i^\vee \rangle + \langle \alpha_{i_3}, \alpha_i^\vee \rangle, \, \forall i \in I.$$

Again $\langle z_3, \alpha_i^\vee \rangle \geq 0$, for $i = i_0, i_1, i_2, i_3$ and z_3 is not dominant if and only if

$$\exists\, i_4 \in I \setminus \{i_0, i_1, i_2, i_3\} \text{ s.t. } \langle \alpha_{i_3}, \alpha_{i_4}^\vee \rangle = -1, \text{ and } \langle z, \alpha_{i_4}^\vee \rangle = 0. \tag{6}$$

(If i_4 as above exists, then it must be unique.)

If (6) holds, then we consider $z_4 := s_{\alpha_{i_4}}(z_3) = z + \alpha_{i_1} + \alpha_{i_2} + \alpha_{i_3} + \alpha_{i_4}$. If z_4 is dominant, then the process of finding w_0 stops here, otherwise, we see that

$$\exists\, i_5 \in I \setminus \{i_0, i_1, i_2, i_3, i_4\} \text{ s.t. } \langle \alpha_{i_4}, \alpha_{i_5}^\vee \rangle = -1, \text{ and } \langle z, \alpha_{i_5}^\vee \rangle = 0. \tag{7}$$

(If i_5 as above exists, then it must be unique.)

Since the number of simple roots is finite, we conclude, using induction, that

$$\exists k \in \{1, \ldots, n-1\} \text{ s.t. } w_0 = s_{\alpha_{i_k}} \circ \ldots \circ s_{\alpha_{i_2}} \circ s_{\alpha_{i_1}},$$

and therefore

$$w_0 z = z + \alpha_{i_1} + \alpha_{i_2} + \ldots + \alpha_{i_k}$$

is dominant. Clearly, by construction:

(a) The α_{i_j}'s appearing in w_0 are distinct;
(b) $\langle z, \alpha_{i_1}^\vee \rangle = -1$;
(c) $\langle z, \alpha_{i_j}^\vee \rangle = 0$ for $j = 2, \ldots, k$;
(d) $\langle \alpha_{i_j}, \alpha_{i_{j+1}}^\vee \rangle = -1$, for $j = 1, \ldots, k-1$.

All that remains to finish the proof of Lemma 3.1 (and therefore Theorem 2.1) is to check that $w_0 z$ satisfies the right-hand side of (1). This is done in the lemma below.

Lemma 3.2. *With notation as above (in particular, $z \in Q(R)$ is assumed to satisfy (2)), suppose that $w_0 z = z + \alpha_{i_1} + \alpha_{i_2} + \ldots + \alpha_{i_k}$ is dominant and conditions* (a)-(d) *hold. Then*

$$\langle w_0 z, \varpi_i \rangle \leq \langle x, \varpi_i \rangle, \, \forall i \in I.$$

Proof. Conditions (2) and (a) together imply that

$$\langle w_0 z, \varpi_i \rangle \leq \langle x, \varpi_i \rangle + 1, \, \forall i \in \{i_1, \ldots, i_k\},$$

and

$$\langle w_0 z, \varpi_i \rangle \leq \langle x, \varpi_i \rangle, \, \forall i \notin \{i_1, \ldots, i_k\}.$$

Therefore, it suffices to check that

$$\langle z, \varpi_i \rangle < \langle x, \varpi_i \rangle, \, \forall i \in \{i_1, \ldots, i_k\},$$

since

$$\langle w_0 z, \varpi_i \rangle = \langle z, \varpi_i \rangle + \sum_{j=1}^{k} \langle \alpha_{i_j}, \varpi_i \rangle.$$

Suppose for a contradiction that $\langle z, \varpi_{i_1} \rangle = \langle x, \varpi_{i_1} \rangle$. Then, since $z \in Q(R)$ satisfies (2), we get that

$$z = x - \sum_{i \in I \setminus \{i_1\}} a_i \, \alpha_i,$$

for some non-negative integers a_i. But x is dominant, so $\langle x, \alpha_{i_1}^\vee \rangle \geq 0$, and also $\langle \alpha_i, \alpha_{i_1}^\vee \rangle \leq 0$, $\forall i \neq i_1$. Thus, since a_i's are non-negative, we have $\langle z, \alpha_{i_1}^\vee \rangle \geq 0$. But this contradicts our assumption (b) and therefore we must have $\langle z, \varpi_{i_1} \rangle < \langle x, \varpi_{i_1} \rangle$.

Assume now that $\langle z, \varpi_{i_m} \rangle = \langle x, \varpi_{i_m} \rangle$, for some $m \in \{2, \ldots, k\}$. Then, for similar reasons to those above, we can write

$$z = x - \sum_{i \in I \setminus \{i_m\}} b_i \, \alpha_i,$$

for some non-negative integers b_i. Since x is dominant, $\langle \alpha_i, \alpha_{i_m}^\vee \rangle \leq 0$, $\forall i \neq i_m$, and b_i's are non-negative, we see that $b_i = 0$, each time $\langle \alpha_i, \alpha_{i_m}^\vee \rangle < 0$, or else (c) would be contradicted. Condition (d) gives $\langle \alpha_{i_{m-1}}, \alpha_{i_m}^\vee \rangle = -1$, so $b_{m-1} = 0$.

We now get that

$$z = x - \sum_{i \in I \setminus \{i_m, i_{m-1}\}} b_i \, \alpha_i,$$

for some non-negative integers b_i. From (d) we have $\langle \alpha_{i_{m-2}}, \alpha_{i_{m-1}}^\vee \rangle = -1$, and so, because x is dominant, b_i's are non-negative, and the non-diagonal entries $\langle \alpha_i, \alpha_j^\vee \rangle$ $(i \neq j)$ of the so-called Cartan matrix are not positive, we conclude, using (c), that $b_{i_{m-2}} = 0$, or equivalently

$$z = x - \sum_{i \in I \setminus \{i_m, i_{m-1}, i_{m-2}\}} b_i \, \alpha_i.$$

We continue this process by induction, to find that

$$z = x - \sum_{i \in I \setminus \{i_m, i_{m-1}, \ldots, i_1\}} b_i \, \alpha_i.$$

But this implies that $\langle z, \varpi_{i_1} \rangle = \langle x, \varpi_{i_1} \rangle$, contradicting the inequality $\langle z, \varpi_i \rangle < \langle x, \varpi_i \rangle$, demonstrated earlier. □

Remark 3.3. As was apparent in the proof of Lemma 1.1, z is always dominant for simply-laced root systems, but it may fail to be dominant for the other root systems. However, when z fails to be dominant, finding a $w_0 \in W$ so that $w_0 z$ is dominant was easier for some root systems than for others. More specifically, if z is not dominant, then $w_0 z = z + \alpha_{i_1} + \alpha_{i_2} + \ldots + \alpha_{i_k}$ for certain distinct α_{i_j}'s, and the number k depends on the root systems. Call *the defect of* z the number k of simple reflections, as in the proof of Lemma 1.1, required to make z dominant. We then call *the defect of the root system* R, denoted defect(R), the maximum of defects of z, for $z \in Q(R)$. With this terminology, we have the following list, which, in a way, tells us the level of difficulty for solving our initial problem, stated in the Lemma 1.1.

1. *defect* $(A_n) = 0$
2. *defect* $(D_n) = 0$
3. *defect* $(E_6) = 0$
4. *defect* $(E_7) = 0$
5. *defect* $(E_8) = 0$
6. *defect* $(B_n) = 1$
7. *defect* $(G_2) = 1$
8. *defect* $(F_4) = 2$
9. *defect* $(C_n) = n - 2$

References

[1] P. BERTHELOT and A. OGUS, *Notes on Crystalline Cohomology*, Princeton Univ. Press, Princeton, NJ, 1978.

[2] M. BRION and S. KUMAR, *Frobenius splitting methods in geometry and representation theory*, Birkhäuser Boston (2004).

[3] N. BOURBAKI, *Elements of Mathematics, Lie Groups and Lie Algebras, Chapters 4-6*, Springer-Verlag, Berlin Heidelberg 2002.

[4] P.-H. CHAUDOUARD and G. LAUMON, *Sur l'homologie des fibres de Springer affines tronqués*, preprint at arXiv:math/0702586

[5] J.-M. FONTAINE and M. RAPOPORT, *Existence de filtrations admissible sur des isocristaux*, Bull. Soc. Math. France **133** (2005), no.1, 73–86.

[6] W. FULTON, *Introduction to Toric Varieties*, Ann. of Math. Stud. 131, Princeton Univ. Press, Princeton, NJ, 1993.

[7] Q. R. GASHI, *Vanishing Results for Toric Varieties Associated to GL_n and G_2*, Transform. Groups, to appear.

[8] R. KOTTWITZ, *On the Hodge-Newton decomposition for split groups*, Int. Math. Res. Not. **2003,** no.26, 1433–1447.

[9] R. KOTTWITZ, *Harmonic analysis on reductive p-adic groups and Lie algebras*, in *Harmonic Analysis, the Trace Formula, and Shimura Varieties*, 393–522, Clay Math. Proc., **4,** Amer. Math. Soc., Providence, RI, 2005.

[10] R. KOTTWITZ and M. RAPOPORT, *On the existence of F-isocrystals*, Comment. Math. Helv. **78** (2003), 153–184.

[11] C. LUCARELLI, *A converse to Mazur's inequality for split classical groups*, Journal of the Inst. Math. Jussieu (2004) **3** (2), 165–183.

[12] B. MAZUR, *Frobenius and the Hodge filtration*, Bull. Amer. Math. Soc. **78** (1972), 653–667.

[13] M. MUSTAŢĂ, *Vanishing theorems on toric varieties*, Tohoku Math. J. (2) **54** (2002), no.3, 451–470.

[14] M. RAPOPORT and M. RICHARTZ, *On the classification and specialization of F-isocrystals with additional structure*, Composito Math. **103** (1996), no.2, 153–181

Department of Mathematics
The University of Chicago,
5734 S. University Avenue,
Chicago, IL 60637
E-mail: qendrim@math.uchicago.edu

ALBANIAN JOURNAL
OF MATHEMATICS
Volume 1, Number 4, Pages 245–252
ISSN 1930-1235: (2007)

HYPERELLIPTIC CURVES WITH a-NUMBER 1 IN SMALL CHARACTERISTIC

ARSEN ELKIN AND RACHEL PRIES

Abstract. For every $g \geq 3$, we show there exists a hyperelliptic curve of genus g with p-rank $g-3$ and a-number 1 in characteristic p when $p=3$ or $p=5$. The method of proof is to show that a generic point of the moduli space of hyperelliptic curves of genus 3 and p-rank 0 has a-number 1. When $p=3$, we also show that this moduli space is irreducible.

1. Introduction

Suppose X is a curve of genus g defined over an algebraically closed field k of characteristic p. The p-torsion of the Jacobian $\mathrm{Jac}(X)$ can be studied using invariants such as the p-rank σ_X and a-number a_X. In Section 2, we define these invariants. Briefly, the *p-rank* of X is $\sigma_X = \dim_{\mathbb{F}_p} \mathrm{Hom}(\mu_p, \mathrm{Jac}(X))$ where the group scheme μ_p is the kernel of Frobenius on $\mathbb{G}_m$. The *a-number* of X is $a_X = \dim_k \mathrm{Hom}(\alpha_p, \mathrm{Jac}(X))$ where the group scheme α_p is the kernel of Frobenius on $\mathbb{G}_a$. It is well known that σ_X, a_X are non-negative integers with $0 \leq \sigma_X + a_X \leq g$.

There are many open problems about the p-rank and a-number of curves. In some sense this is surprising, since there are algorithms to compute the p-rank and the a-number of X for a given prime p and a given curve X. However, these algorithms are not well-suited for proving existence results for curves of arbitrary genus in arbitrary characteristic. For this reason, many of the existence results on this topic are non-constructive and rely on deep theorems from arithmetic geometry, e.g., [2, Thm. 2.3].

A result from [8] is that, for every prime p and every $g \geq 3$, there exists a k-curve X of genus g with p-rank $g-3$ and a-number 1. The author also gives a strategy for extending this result to the case of hyperelliptic curves and explains some of the difficulties involved with this strategy. In this paper, we carry out this strategy when $p=3$ and $p=5$, which yields the following result (found in Section 4).

Corollary 1.1. *Suppose $g \geq 3$. Let $p=3$ or $p=5$. Then there exists a hyperelliptic curve of genus g in characteristic p with p-rank $g-3$ and a-number 1.*

For the proof, we consider the moduli space $\mathcal{H}_3 \cap V_{3,0}$ whose points correspond to hyperelliptic curves of genus 3 with p-rank 0. When $p=3$ and $p=5$, we give an explicit proof that every generic point of this moduli space has a-number 1 in Section 3. This provides the base case of an inductive process found in [8]. Using induction on g, we conclude that the locus of curves having a-number 1 is an open and dense subspace of the moduli space of hyperelliptic curves of genus g and p-rank $g-3$ (Theorem 4.2).

We also show that $\mathcal{H}_3 \cap V_{3,0}$ is irreducible when $p = 3$ (Proposition 3.5). It is an open question whether $\mathcal{H}_3 \cap V_0$ is irreducible when $p > 3$. We describe the computational complexity of this problem in Section 3.4.

The second author was partially supported by NSF grant DMS-07-01303. We thank J. Achter for his comments on drafts of this paper.

2. Invariants of the p-torsion of Jacobians

2.1. **The p-rank and a-number.** Throughout the paper, we work over an algebraically closed field k of characteristic p. The group scheme μ_p is the kernel of Frobenius on $\mathbb{G}_m$ and the group scheme α_p is the kernel of Frobenius on $\mathbb{G}_a$. As schemes, $\mu_p \simeq \operatorname{Spec}(k[x]/(x-1)^p)$ and $\alpha_p \simeq \operatorname{Spec}(k[x]/x^p)$. See [4, A.3] for more details about these group schemes.

Suppose X is a smooth projective k-curve of genus g with Jacobian $\operatorname{Jac}(X)$. The *p-rank* of X is $\sigma_X = \dim_{\mathbb{F}_p} \operatorname{Hom}(\mu_p, \operatorname{Jac}(X))$ and the *a-number* of X is $a_X = \dim_k \operatorname{Hom}(\alpha_p, \operatorname{Jac}(X))$. The p-rank is the integer σ_X such that the number of p-torsion points of $\operatorname{Jac}(X)$ is p^{σ_X}. It is well-known that $0 \leq \sigma_X + a_X \leq g$.

2.2. **Moduli spaces of curves with given invariants.** Let $\mathcal{M}_g$ denote the moduli space of smooth projective curves of genus g defined over k. Let $\mathcal{H}_g \subset \mathcal{M}_g$ denote the sublocus consisting of hyperelliptic curves. The dimension of $\mathcal{H}_g$ is $2g-1$. Let $V_{g,\sigma} \subset \mathcal{M}_g$ denote the closed sublocus consisting of curves of genus g with p-rank at most σ. Every irreducible component of $\mathcal{H}_g \cap V_{g,\sigma}$ has dimension $g - 1 + \sigma$ by [3, Thm. 1].

Let $T_{g,2} \subset \mathcal{M}_g$ denote the closed sublocus of curves with a-number at least 2. Recall that $T_{g,2} \subset V_{g,g-2}$. If $g \geq 2$ and $\sigma = g-2$, the generic point of every irreducible component of $\mathcal{H}_g \cap V_{g,g-2}$ has a-number 1, [8, Thm. 4.1]. It follows that $\dim(\mathcal{H}_g \cap T_{g,2}) \leq 2g-4$. In particular, $\dim(\mathcal{H}_3 \cap T_{3,2}) \leq 2$.

Remark 2.1. When $p = 2$, every hyperelliptic cover is wildly ramified. As a result, the computation of the p-rank or a-number of a hyperelliptic curve differs significantly when $p = 2$ from the case when p is odd. For example, every hyperelliptic curve of genus 3 and p-rank 0 has a-number 2 [2, 3.2]. For every finite field $\mathbb{F}$ of characteristic 2, and for $0 \leq \sigma \leq 3$, there is a formula for the number of isomorphism classes of hyperelliptic curves defined over $\mathbb{F}$ with genus 3 and p-rank σ [5, Table 3].

2.3. **Computing the p-rank and a-number.** Suppose that $p \geq 3$ and that X is hyperelliptic. There is a $\mathbb{Z}/2$-Galois cover $\phi : X \to \mathbb{P}^1_k$ with $2g+2$ distinct branch points. Without loss of generality, we suppose ϕ is branched at ∞ and choose an affine equation for ϕ of the form $y^2 = f(x)$, where $f(x) \in k[x]$ is a polynomial of degree $2g+1$.

Let c_s denote the coefficient of x^s in the expansion of $f(x)^{(p-1)/2}$. For $0 \leq \ell \leq g-1$, let A_ℓ be the $g \times g$ matrix whose ijth entry is $(c_{ip-j})^{p^\ell}$. The matrix A_0 is the Hasse-Witt matrix of X. The Cartier-Manin matrix is $M = (\prod_{\ell=0}^{g-1} A_\ell)$.

Lemma 2.2. *Suppose X is a hyperelliptic curve of genus g with equation $y^2 = f(x)$ as above.*

1. *The a-number of X is $a_X = g - r$ where r is the rank of A_0.*
2. *The p-rank of X is $\sigma_X = \operatorname{rank}(M)$.*

Proof. The Hasse-Witt matrix of X is the matrix for the action of Frobenius on $H^1(X, \mathcal{O}_X)$. By duality, one can consider the matrix of the Cartier operator on $H^0(X, \Omega^1_X)$ instead. The result then follows from [11]. □

Remark 2.3. It is a general phenomenon that $a_X = 0$ occurs only when $\sigma_X = g$ [6, p.416]. Lemma 2.2 illustrates this for hyperelliptic curves: if $a_X = 0$ then A_0 is invertible, and thus M is invertible, which implies that $\sigma_X = g$.

3. Hyperelliptic curves of genus 3

3.1. **Parametrization of hyperelliptic curves of genus 3.** Let Y be a smooth hyperelliptic curve of genus 3. Then Y has an affine equation $y^2 = f(x)$ where $f(x) \in k[x]$ has distinct roots and is of degree 7. We say that the equation $y^2 = f(x)$ is in *standard form* if $f(x) = x^7 + ax^6 + bx^5 + cx^4 + dx^3 + ex^2 + x$ for some $a, b, c, d, e \in k$.

Lemma 3.1. *Every smooth hyperelliptic curve Y of genus 3 has an affine equation $y^2 = f(x)$ in standard form. There are only finitely many choices of $f(x)$ so that $y^2 = f(x)$ is an affine equation in standard form for Y.*

Proof. If Y is hyperelliptic then there is a morphism $\phi : Y \to \mathbb{P}^1_k$ of degree 2. If Y has genus 3 then the Riemann-Hurwitz formula implies that the branch locus B of ϕ contains exactly 8 points. After a change of coordinates on $\mathbb{P}^1_k$, we can suppose $0, \infty \in B$. Then ϕ is given by an affine equation of the form $y^2 = f(x)$ for some $f(x) \in k[x]$ with $\deg(f(x)) = 7$ and $f(0) = 0$. Write $f(x) = \sum_{i=1}^{7} a_i x^i$ where $a_i \in k$ and $a_1 a_7 \neq 0$.

Consider a change of coordinates $T(y) = \alpha y$ and $T(x) = \beta x$ with $\alpha, \beta \in k$ and $\alpha\beta \neq 0$. Let $f_T(x) = (a_7\beta^7/\alpha^2)x^7 + \cdots + (a_1\beta/\alpha^2)x$. Then $y^2 = f_T(x)$ is another affine equation for Y. Let $\alpha, \beta \in k^*$ be solutions to $\alpha = (a_1^7/a_7)^{1/12}$ and $\beta = (a_1/a_7)^{1/6}$. Then the equation $y^2 = f_T(x)$ is in standard form.

Suppose $y^2 = f_1(x)$ and $y^2 = f_2(x)$ are two equations for Y in standard form. Then there is a change of coordinates $T : k[x, y]/(y^2 - f_1(x)) \to k[x, y]/(y^2 - f_2(x))$. Since the hyperelliptic involution is in the center of $\mathrm{Aut}(Y)$, the change of coordinates descends to an automorphism T of $\mathbb{P}^1_k$. Also T stabilizes $\{0, \infty\}$. After possibly composing T with an inversion $x \mapsto 1/x$, we can suppose T fixes 0 and ∞. It follows that $T(x) = \beta x$ and $T(y) = \alpha y$ for some $\alpha, \beta \in k^*$. Then $\beta^7/\alpha^2 = \beta/\alpha^2 = 1$. Thus $\beta^6 = 1$ so there are at most 6 choices for β and for each of these there are at most 2 choices for α. □

3.2. **Irreducibility of $\mathcal{H}_3 \cap V_{3,0}$ when $p = 3$.** In this section, suppose $p = 3$. The main result of the section is that $\mathcal{H}_3 \cap V_{3,0}$ is irreducible. In the next lemma, we first show that all smooth hyperelliptic curves Y of genus 3 have a-number at most 1. The lemma is a special case of [1, Thm. 1], but we include a proof for the convenience of the reader. See [10] for similar results for curves that are not hyperelliptic.

Lemma 3.2. *If $p = 3$, then $\mathcal{H}_3 \cap T_{3,2} = \emptyset$. In other words, there are no smooth hyperelliptic curves of genus 3 with a-number at least 2.*

Proof. Suppose Y is a smooth hyperelliptic curve of genus 3. By Lemma 3.1, Y has an affine equation $y^2 = f(x)$ where $f(x) = x^7 + ax^6 + bx^5 + cx^4 + dx^3 + ex^2 + x$. If $p = 3$, the entries of A_0 are given by the coefficients of $f(x)$:

$$A_0 = \begin{pmatrix} e & 1 & 0 \\ b & c & d \\ 0 & 1 & a \end{pmatrix}.$$

If $a_Y \geq 2$, then $\operatorname{rank}(A_0) \leq 1$ by Lemma 2.2(1). This implies $e = b = d = a = 0$. Then $f(x) = x(x^2 + c^{1/3}x + 1)^3$ does not have distinct roots which contradicts the hypothesis that Y is smooth. Thus $a_Y \leq 1$. □

By Lemma 3.2, every point of the two-dimensional space $\mathcal{H}_3 \cap V_{3,0}$ has a-number 1 when $p = 3$. In fact, we can say more about the geometry of $\mathcal{H}_3 \cap V_{3,0}$ when $p = 3$. The next result gives necessary and sufficient conditions on the five parameters $a, \ldots, e$ for Y to have p-rank 0.

Lemma 3.3. *Suppose Y is a smooth hyperelliptic curve with affine equation $y^2 = f(x)$ where $f(x) = x^7 + ax^6 + bx^5 + cx^4 + dx^3 + ex^2 + x$. Then Y has p-rank 0 in exactly the following cases:*

(1) *$d = 0$, $a = b + c^4 = e + c^3 = 0$;*
(2) *$d \neq 0$, $b^3 + c^{12} + c^9a + d^3 + a^4 = d^6e + d^6c^3 + a^9 = d^9c^3 + d^9a + d^3a^9 + a^{13} = 0$.*

Proof. By Lemma 2.2(2), Y has p-rank 0 exactly when $M = A_0A_1A_2$ is the zero matrix. One computes that the matrix M has entries m_{ij} where:

$$\begin{aligned}
m_{11} &= e^{13} + b^3e^9 + b^9e + b^9c^3; \\
m_{12} &= e^4 + b^3 + c^9e + c^{12} + d^3; \\
m_{13} &= d^9e + d^9c^3 + d^3a^9; \\
m_{21} &= e^{12}b + e^9cb^3 + b^{10} + b^9c^4 + b^9d; \\
m_{22} &= be^3 + cb^3 + c^9b + c^{13} + c^9d + cd^3 + da^3; \\
m_{23} &= d^9b + d^9c^4 + d^{10} + a^9cd^3 + a^{12}d; \\
m_{31} &= b^3e^9 + b^9c^3 + b^9a; \\
m_{32} &= b^3 + c^{12} + c^9a + d^3 + a^4; \\
m_{33} &= d^9c^3 + d^9a + d^3a^9 + a^{13}.
\end{aligned}$$

Let $I \subset k[a, b, c, d, e]$ be the ideal $I = (m_{ij} \mid 1 \leq i, j \leq 3)$. Consider a point $w = (a, b, c, d, e) \in \mathbb{A}^5_k$. Let $V(I) \subset \mathbb{A}^5_k$ be the variety of I. Then Y has p-rank 0 if and only if $w \in V(I)$.

(1) Suppose $w \in V(I)$ and $d = 0$. Then equation m_{33} implies $a = 0$. Then equation m_{32} implies $b + c^4 = 0$. If $e = 0$, then equation m_{11} implies $c = 0$ and so $e + c^3 = 0$. (Note that $y^2 = x^7 + x$ is not smooth, so the case $e = 0$ can be disregarded anyway.) If $e \neq 0$, then equation m_{12} implies $e + c^3 = 0$. Conversely, if $d = a = b + c^4 = e + c^3 = 0$, then a computer calculation shows that $w \in V(I)$.
(2) Suppose $w \in V(I)$ and $d \neq 0$, then equation m_{13} implies $d^6e + d^6c^3 + a^9 = 0$. Also equation m_{32} implies $b^3 + c^{12} + c^9a + d^3 + a^4 = 0$. Then equation m_{33} implies $d^9c^3 + d^9a + d^3a^9 + a^{13} = 0$. Conversely, after solving for e, b, and then c, and substituting them into m_{ij}, a computer calculation shows that $w \in V(I)$.

□

Lemma 3.4. *Let $I \subset k[a,b,c,d,e]$ be the ideal $I = (m_{ij} \mid 1 \leq i,j \leq 3)$ as above. Then $V(I)$ is irreducible with dimension two.*

Proof. Suppose that $(a,b,c,d,e) \in V(I)$ with $d \neq 0$. Using the equations from Lemma 3.3(2), one can solve for b and e in terms of $a^{1/3}, c, d$, and then one can solve for c in terms of $a^{1/3}$ and d. Namely, $e = 2c^3 + 2a^9/d^6$ and $b = 2c^4 + 2c^3a^{1/3} + 2d + 2a^{4/3}$. Also $c = 2a^{1/3} + 2a^3/d^2 + 2a^{13/3}/d^3$. Thus there are formulae $b(a^{1/3}, d)$, $c(a^{1/3}, d)$, $e(a^{1/3}, d)$ for b, c, e in terms of $a^{1/3}, d$.

Let $S = \mathrm{Spec}(k[a^{1/3}, d, d^{-1}])$. Note that S is irreducible with dimension 2. Let $C \subset \mathbb{A}^5_k$ be the closed subspace of points (a,b,c,d,e) with $d = 0$. Let $U = \mathbb{A}^5 - C$. The morphism $G((a^{1/3}, d)) = (a, b(a^{1/3}, d), c(a^{1/3}, d), d, e(a^{1/3}, d))$ yields an isomorphism $G : S \to V(I) \cap U$. Thus $V(I) \cap U$ is irreducible with dimension two.

It remains to show that $V(I) \cap C$ is in the boundary of $V(I) \cap U$. Let $W \subset V(I) \cap U$ be the closed locus where $d^2 + a^2 = 0$. Recall that if $w \in V(I)$ then $d^6e + d^6c^3 + a^9 = 0$ and $(b + c^4)^3 + c^9a + d^3 + a^4 = 0$. If also $w \in W$, then $e + c^3 + a^3 = 0$. When $a = d = 0$, these relations imply that $e + c^3 = b + c^4 = 0$. Thus every point of $V(I) \cap C$ is in the boundary of $V(I) \cap U$. □

Proposition 3.5. *When $p = 3$, the moduli space $\mathcal{H}_3 \cap V_{3,0}$ is irreducible.*

Proof. Let $\Delta \subset \mathbb{A}^5_k$ be the closed subset of all (a,b,c,d,e) so that $f(x)$ has multiple roots. Let $U' = \mathbb{A}^5_k - \Delta$. There is a morphism $\tau : U' \to \mathcal{H}_3$ which is surjective (and finite-to-one) by Lemma 3.1. Then $\tau^{-1}(\mathcal{H}_3 \cap V_{3,0}) = V(I) \cap U'$. By Lemma 3.4, $V(I)$ is irreducible. Thus $\mathcal{H}_3 \cap V_{3,0}$ is irreducible when $p = 3$. □

3.3. **The case when $p = 5$.** In this section, suppose $p = 5$. The computations of the previous section become more elaborate. We show that $\mathcal{H}_3 \cap T_{3,2}$ has exactly one irreducible component of dimension two and that its generic point has p-rank 1. Using this, we show that the generic point of every irreducible component of $\mathcal{H}_3 \cap V_{3,0}$ has a-number 1.

Lemma 3.6. *If $p = 5$, then $\mathcal{H}_3 \cap T_{3,2}$ contains exactly one irreducible component of dimension 2 and the generic point of this component has a-number 2 and p-rank 1.*

Proof. If $p = 5$, the entries of A_0 are given by some of the coefficients of $f(x)^2$:

$$A_0 = \begin{pmatrix} 2d + e^2 & 2e & 1 \\ 2e + 2ad + 2bc & 2 + 2ae + c^2 + 2bd & 2cd + 2a + 2be \\ 1 & 2a & 2b + a^2 \end{pmatrix}.$$

If Y has a-number at least 2, then the rank of A_0 is at most 1. Thus the first two rows of A_0 are a non-zero scalar multiple of the third row. This implies that $(a,b,c,d,e) \in V(J)$ where $J \subset k[a,b,c,d,e]$ is the ideal (t_1, t_2, t_3, t_4) where:

$$\begin{aligned} t_1 &= 4ad + 2ae^2 + 3e; \\ t_2 &= 4bd + 2be^2 + 2a^2d + a^2e^2 + 4; \\ t_3 &= 2ae + 4a^2d + 4abc + 3 + 4c^2 + 3bd; \\ t_4 &= 2be + 4bad + 4b^2c + 2a^2e + 2a^3d + 2a^2bc + 3cd + 3a. \end{aligned}$$

By equation t_1, if $a = 0$ then $e = 0$. Then $bd = -1$ and $c = 0$. Similarly, if $e = 0$, then $ad = 0$ and $bd = -1$, which gives $a = c = 0$. In either case, this yields a component of $V(J)$ of dimension 1.

Suppose $ae \neq 0$. Then equation t_1 implies that $d = 2e^2 + 3e/a$. After making this substitution, equation t_2 implies that $b = 2a^2 + 3a/e$. After making this substitution, equations t_3 and t_4 simplify as follows:

$$t_3' = 3a^3ce + 2a^2c + 4c^2e;$$
$$t_4' = 4ca^4e + ca^3 + ce^2a + 4ce^3.$$

If $c \neq 0$, then one can show that $c = 3a^3 + 2a^2/e$. Also $c \neq 0$ implies $ae \neq 1$. Another computation then shows that there is a relation between a and e, namely $a^3 = e^3$. Thus the intersection $V(J) \cap \{c \neq 0\}$ has dimension one, which yields a subset of $\mathcal{H}_3 \cap T_{3,2}$ having dimension one.

Otherwise, if $c = 0$, then $t_3' = t_4' = 0$. In other words,

$$V(J) \cap \{ae \neq 0,\ c = 0\} = \{(a, b, 0, d, e) \mid b = 2a^2 + 3a/e,\ d = 2e^2 + 3e/a\}.$$

Thus there is a unique irreducible component of $V(J)$ having dimension two. As in the proof of Proposition 3.5, there is a surjective finite-to-one morphism $\tau : U' \to \mathcal{H}_3$. Then $\tau^{-1}(\mathcal{H}_3 \cap T_{3,2}) = V(J) \cap U'$. This yields a unique irreducible component η of $\mathcal{H}_3 \cap T_{3,2}$ having dimension two.

We now find a point $w \in V(J)$ with $ae \neq 0$ and $c = 0$ so that the corresponding curve Y_w is a smooth hyperelliptic curve of genus 3 with a-number 2 and p-rank 1. Let $\gamma \in \mathbb{F}_{25}$ be a root of x^2-2. Consider the point $w = (\gamma, 4+3\gamma, 0, 2+4\gamma, 1) \in V(J)$. One can compute that the discriminant of $f(x) = x^7 + \gamma x^6 + (4 + 3\gamma)x^5 + (2 + 4\gamma)x^3 + x^2 + x$ is 4 and so $f(x)$ has distinct roots. Thus Y_w is a smooth hyperelliptic curve of genus 3. Also Y_w has a-number 2 since $w \in V(J)$. One can compute that

$$A_0(w) = \begin{pmatrix} 3\gamma & 2 & 1 \\ 4\gamma + 3 & 1 + \gamma & 3 + 3\gamma \\ 1 & 2\gamma & \gamma \end{pmatrix}.$$

Thus $M = A_0A_1A_2$ simplifies to:

$$M(w) = \begin{pmatrix} 2\gamma & 3 & 4 \\ \gamma + 2 & 4\gamma + 4 & 2\gamma + 2 \\ 4 & 3\gamma & 4\gamma \end{pmatrix}.$$

Then $Y_w \in \eta$ has p-rank 1 since $\operatorname{rank}(M(w)) = 1$. The p-rank can only decrease under specialization. Thus the generic point of η has p-rank 1 and a-number 2. □

3.4. **Complexity Analysis.** As the characteristic increases, the sort of analysis on the a-number and p-rank performed in previous sections becomes prohibitively complicated. To see this, let $f(x) = a_7x^7+a_6x^6+a_5x^5+a_4x^4+a_3x^3+a_2x^2+a_1x+a_0$. Then every coefficient of $g(x) = f(x)^{(p-1)/2}$ is homogeneous of degree $(p-1)/2$ when considered as a polynomial in $k[a_0, \ldots, a_7]$.

The coefficient of x^{p-1} in $g(x)$ contains a monomial $a_2^{(p-1)/2}$. The degree in a_2 of any other coefficient of $g(x)$ is strictly less than $(p-1)/2$. Thus the entry a_{11} of $A_0 = (a_{ij})$ contains a monomial $a_2^{(p-1)/2}$, and all other entries of this matrix have smaller degree as polynomials in a_2. The non-homogeneous version of this statement is that the highest power of e appearing in the Cartier-Manin matrix for the curve $y^2 = x^7 + ax^6 + bx^5 + cx^4 + dx^3 + ex^2 + x$ appears in the monomial $e^{(p-1)/2}$ in the entry a_{11}.

This, in turn, implies that a_{11}^p contains a monomial $e^{p(p-1)/2}$, and this is the highest degree to which e appears in the entries of the matrix $A_1 = (a_{ij}^p)$. Similarly, $a_{11}^{p^2}$ contains $e^{p^2(p-1)/2}$, and the degree of e is smaller in all the other entries of $A_2 = (a_{ij}^{p^2})$. Therefore, $e^{(1+p+p^2)(p-1)/2} = e^{(p^3-1)/2}$ is a monomial in the entry m_{11} of the product $(m_{ij}) := A_0A_1A_2$. A similar analysis can be performed for a_4 (or c) in the entry $m_{2,2}$ and for a_6 (or a) in the entry $m_{3,3}$.

This discussion demonstrates that the entries of the matrix A_0, whose rank is analyzed in connection with the a-number, contain monomials of degree $(p-1)/2$ in a, c, and e. The entries of the matrix $A_0A_1A_2$, examined for p-rank, have the same variables appearing with degrees $(p^3-1)/2$. In particular, the locus $\mathcal{H}_3 \cap V_{3,0}$ corresponds to the vanishing of nine equations in five variables, at least three of which have degree $(p^3-1)/2$ in some variable. The difficulty of analyzing these two invariants grows accordingly.

4. Application: hyperelliptic curves with p-rank $g-3$ and a-number 1

Let $g \geq 3$. Suppose X is a curve of genus g with p-rank $g-3$. By Remark 2.3, there are three possibilities for the a-number of X, namely $a_X \in \{1,2,3\}$.

Remark 4.1. If X has genus g and p-rank $g-3$, there are four possibilities for the isomorphism class of the group scheme $\mathrm{Jac}(X)[p]$. Of these, there is a unique group scheme with p-rank $g-3$ and a-number 1. It is of the form $(\mathbb{Z}/p \oplus \mu_p)^{g-3} \oplus I_{3,1}$ where $I_{3,1}$ is the unique group scheme of rank 6, p-rank 0, and a-number 1. The covariant Dieudonné module for $I_{3,1}$ is $E/E(F^3 - V^3)$ where $E = k[F,V]$ is a non-commutative ring generated by Frobenius and Verschiebung [9, Lemma 3.1].

Theorem 4.2. *Suppose $g \geq 3$. Let $p = 3$ or $p = 5$. Then the generic point of every irreducible component of $\mathcal{H}_g \cap V_{g,g-3}$ has a-number 1.*

Proof. The proof is by induction on g with base case $g = 3$. For $p = 3$, $\mathcal{H}_3 \cap T_{3,2} = \emptyset$ by Lemma 3.2. Thus every point of $\mathcal{H}_3 \cap V_{3,0}$ has a-number 1. For $p = 5$, there is a unique irreducible component η of $\mathcal{H}_3 \cap T_{3,2}$ with dimension 2 and its generic point has p-rank 1 by Lemma 3.6. Every irreducible component ξ of $\mathcal{H}_3 \cap V_{3,0}$ has dimension 2 and has generic point with p-rank 0. Thus $\xi \subsetneq T_{3,2}$. So the generic point of every irreducible component of $\mathcal{H}_3 \cap V_{3,0}$ has a-number 1.

For $g \geq 4$, the result follows immediately from [8, Prop. 3.6]. Here is the basic idea of the inductive proof. The compactification $\overline{\mathcal{M}}_g$ of $\mathcal{M}_g$ contains a boundary component Δ_0 whose generic point is a singular curve Z which self-intersects in an ordinary double point. The normalization $\tilde{Z}$ of Z is a smooth curve of genus $g-1$. The p-rank of $\tilde{Z}$ is $\sigma_{\tilde{Z}} = \sigma_Z - 1$. One proves that the closure in $\overline{\mathcal{M}}_g$ of each component of $\mathcal{H}_g \cap V_{g,g-3}$ intersects Δ_0. Then the proof relies on a dimension count for components of Δ_0 that satisfy certain conditions on the p-rank and a-number. □

Corollary 4.3. *Suppose $g \geq 3$. Let $p = 3$ or $p = 5$. There is a family of dimension $2g-4$ consisting of smooth hyperelliptic curves of genus g with p-rank $g-3$ and a-number 1.*

Proof. By Theorem 4.2, the locus of smooth hyperelliptic curves of genus g with p-rank $g-3$ and a-number 1 is open (and dense) in $\mathcal{H}_g \cap V_{g,g-3}$. The result follows since $\dim(\mathcal{H}_g \cap V_{g,g-3}) = 2g-4$ by [3, Thm. 1]. □

Remark 4.4. Here are two strategies for extending Theorem 4.2 to larger characteristic.

(1) By [7, 5.12(4)], for all $p \geq 3$, there exists a hyperelliptic curve of genus 3 with a-number 1. The first strategy would be to see if $\mathcal{H}_3 \cap V_{3,0}$ is irreducible for all $p \geq 3$. If so, the generic point of $\mathcal{H}_3 \cap V_{3,0}$ would have a-number 1 and the result would follow from [8, Prop. 3.6].

(2) By [3, Cor. 4], for all $p \geq 5$, there exists a hyperelliptic curve of genus 3 with a-number 2 and p-rank 1. The second strategy would be to prove that every irreducible component of $\mathcal{H}_3 \cap T_{g,2}$ of dimension two contains a point with p-rank 1. Then the generic point of every irreducible component of $\mathcal{H}_3 \cap V_{3,0}$ would have a-number 1 and the result would again follow from [8, Prop. 3.6].

References

[1] A. Elkin. The rank of the Cartier operator on cyclic covers of the projective line. to appear in J. Algebra, mathAG/0708.0431.

[2] C. Faber and G. van der Geer. Complete subvarieties of moduli spaces and the Prym map. *J. Reine Angew. Math.*, 573:117–137, 2004. arXiv:math.AG/0305334.

[3] D. Glass and R. Pries. Hyperelliptic curves with prescribed p-torsion. *Manuscripta Math.*, 117(3):299–317, 2005. arXiv:math.NT/0401008.

[4] E. Goren. *Lectures on Hilbert modular varieties and modular forms*, volume 14 of *CRM Monograph Series*. American Mathematical Society, Providence, RI, 2002. With the assistance of Marc-Hubert Nicole.

[5] E. Nart and D. Sadornil. Hyperelliptic curves of genus three over finite fields of even characteristic. *Finite Fields Appl.*, 10(2):198–220, 2004.

[6] P. Norman and F. Oort. Moduli of abelian varieties. *Ann. of Math. (2)*, 112(3):413–439, 1980.

[7] F. Oort. Hyperelliptic supersingular curves. In *Arithmetic algebraic geometry (Texel, 1989)*, volume 89 of *Progr. Math.*, pages 247–284. Birkhäuser Boston, Boston, MA, 1991.

[8] R. Pries. The p-torsion of curves with large p-rank. to appear in International Journal of Number Theory, math.AG/0601596.

[9] R. Pries. A short guide to p-torsion of abelian varieties in characteristic p. to appear in Computational Arithmetic Geometry, CONM, AMS.

[10] R. Re. The rank of the Cartier operator and linear systems on curves. *J. Algebra*, 236(1):80–92, 2001.

[11] N. Yui. On the Jacobian varieties of hyperelliptic curves over fields of characteristic $p > 2$. *J. Algebra*, 52(2):378–410, 1978.

Arsen Elkin,
Colorado State University,
Fort Collins, CO, 80521,
elkin@math.colostate.edu.

Rachel Pries,
Colorado State University,
Fort Collins, CO, 80521,
pries@math.colostate.edu.

ALBANIAN JOURNAL
OF MATHEMATICS
Volume 1, Number 4, Pages 253–270
ISSN 1930-1235: (2007)

THETANULLS OF CYCLIC CURVES OF SMALL GENUS

E. PREVIATO, T. SHASKA, AND G. S. WIJESIRI

Abstract. We study relations among the classical thetanulls of cyclic curves, namely curves $\mathcal{X}$ (of genus $g(\mathcal{X}) > 1$) with an automorphism σ such that σ generates a normal subgroup of the group G of automorphisms, and $g(\mathcal{X}/\langle\sigma\rangle) = 0$. Relations between thetanulls and branch points of the projection are the object of much classical work, especially for hyperelliptic curves, and of recent work, in the cyclic case. We determine the curves of genus 2 and 3 in the locus $\mathcal{M}_g(G, \mathbf{C})$ for all G that have a normal subgroup $\langle\sigma\rangle$ as above, and all possible signatures $\mathbf{C}$, via relations among their thetanulls.

1. Introduction

In this paper we consider cyclic algebraic curves, over the complex numbers. These are by definition compact Riemann surfaces $\mathcal{X}$ of genus $g > 1$ (unless we allow singular points, as noted below, so as not attach unnecessary qualifications to a definition or statement), admitting an automorphism σ such that $\mathcal{X}/\sigma \cong \mathbb{P}^1$ and σ generates a normal subgroup of the automorphism group $Aut(\mathcal{X})$ of $\mathcal{X}$. When the curve is hyperelliptic, we insist that the curve have "extra automorphisms", in particular σ is not the hyperelliptic involution. Note that the condition implies to having an equation $y^n = f(x)$ for the curve, where x is an affine coordinate on $\mathbb{P}^1$, σ has order n, and $1, y, \sigma y, ..., \sigma^{n-1}y$ is a basis of $\mathbb{C}(\mathcal{X})/\mathbb{C}(x)$. Naturally, the branch points of $\pi : \mathcal{X} \to \mathbb{P}^1$, together with the signature $\mathbf{C}$ of the cover (its monodromy up to conjugation) provide algebraic coordinates for the curve in moduli, though the same curve could be represented in different ways. The problem of expressing these algebraic data in terms of the transcendental (period matrix, thetanulls, e.g.) is classical. We use below formulas for genus-2 curves due to Rosenhein and Picard, Thomae's formulas for hyperelliptic curves, and a recent generalization of the latter for cyclic curves with $\langle s\rangle \cong C_3$, where we denote by C_n the cyclic group of order n, due to Nakayashiki [8]; several other authors recently obtained partial generalizations to cyclic curves also. We do not aim here at a complete account of the classical or contemporary work on these problems.

Cyclic curves are rare in the moduli space $\mathcal{M}_g$ of smooth curves, and it is desirable to characterize their locus, by algebraic conditions on the equation of the curve, or by analytic conditions on its Abelian coordinates, in other words, theta functions, and better yet, by both. We achieve this for genera 2 and 3, making

2000 *Mathematics Subject Classification.* 14H32, 14H37, 14K25.

Key words and phrases. Theta functions, algebraic curves, moduli spaces, automorphism groups.

⋆ The first author was supported in part by the NSF grant NSF-DMS-0205643.

⋆⋆ The second author was partially supported by an NSF grant and by the NATO grant ICS.EAP.ASI. No. 982903.

recourse to classical formulas, some recent results of Hurwitz space theory, and symbolic manipulation.

The contents of the paper are as follows. In section 2 we recall the notation for Riemann's theta function, as well as classical facts on theta characteristics; we recall Frobenius' and Thomae's formulas for hyperelliptic curves. In sections 3 and 4, respectively, we specialize to the case of genera 2 and 3, we recall recent results on $\mathcal{M}_g(G, \mathbf{C})$, and we calculate thetanull constraints that define the loci of the cyclic curves, using the results we cited. The cleanest case is the one of genus 2 and $\langle\sigma\rangle \cong C_2$, which was classified by Jacobi who gave a condition in terms the branch points of the hyperelliptic involution; such a condition was extended, in principle, to any curve in $\mathcal{M}_g(C_n, \mathbf{C})$, cf. [3] or [9], but the algebraic equation satisfied by the branch points would rapidly become intractable with the size of n.

2. Preliminaries

In this section we give a brief description of the basic setup. All of this material can be found in any standard book on theta functions.

Let $\mathcal{X}$ be a genus $g \geq 2$ algebraic curve. We choose a symplectic homology basis for $\mathcal{X}$, say $\{A_1, \ldots, A_g, B_1, \ldots, B_g\}$, such that the intersection products $A_i \cdot A_j = B_i \cdot B_j = 0$ and $A_i \cdot B_j = \delta_{ij}$, where δ_{ij} is the Kronecker delta. We choose a basis $\{w_i\}$ for the space of holomorphic 1-forms such that $\int_{A_i} w_j = \delta_{ij}$. The matrix $\Omega = \left[\int_{B_i} w_j\right]$ is the *period matrix* of $\mathcal{X}$. The columns of the matrix $[I \ |\Omega]$ form a lattice L in $\mathbb{C}^g$ and the Jacobian of $\mathcal{X}$ is Jac $(\mathcal{X}) = \mathbb{C}^g/L$. Let $\mathcal{H}_g$ be the *Siegel upper-half space*. Then $\Omega \in \mathcal{H}_g$ and there is an injection

$$\mathcal{M}_g \hookrightarrow \mathcal{H}_g/Sp_{2g}(\mathbb{Z}) =: \mathcal{A}_g$$

where $Sp_{2g}(\mathbb{Z})$ is the *symplectic group*. For any $z \in \mathbb{C}^g$ and $\tau \in \mathcal{H}_g$ *Riemann's theta function* is defined as

$$\theta(z, \tau) = \sum_{u \in \mathbb{Z}^g} e^{\pi i(u^t \tau u + 2u^t z)}$$

where u and z are $g-$dimensional column vectors and the products involved in the formula are matrix products. The fact that the imaginary part of τ is positive makes the series absolutely convergent over any compact sets. Therefore, the function is analytic. The theta function is holomorphic on $\mathbb{C}^g \times \mathcal{H}_g$ and satisfies

$$\theta(z + u, \tau) = \theta(z, \tau), \quad \theta(z + u\tau, \tau) = e^{-\pi i(u^t \tau u + 2z^t u)} \cdot \theta(z, \tau),$$

where $u \in \mathbb{Z}^g$; see [6] for details. Any point $e \in$ Jac $(\mathcal{X})$ can be written uniquely as $e = (b, a)\begin{pmatrix} 1_g \\ \Omega \end{pmatrix}$, where $a, b \in \mathbb{R}^g$. We shall use the notation $[e] = \begin{bmatrix} a \\ b \end{bmatrix}$ for the characteristic of e. For any $a, b \in \mathbb{Q}^g$, the theta function with rational characteristics is defined as

$$\theta \begin{bmatrix} a \\ b \end{bmatrix}(z, \tau) = \sum_{u \in \mathbb{Z}^g} e^{\pi i((u+a)^t \tau (u+a) + 2(u+a)^t (z+b))}.$$

When the entries of column vectors a and b are from the set $\{0, \frac{1}{2}\}$, then the characteristics $\begin{bmatrix} a \\ b \end{bmatrix}$ are called the *half-integer characteristics*. The corresponding theta functions with rational characteristics are called *theta characteristics*. A scalar obtained by evaluating a theta characteristic at $z = 0$ is called a *theta*

constant. Points of order n on Jac $\mathcal{X}$ are called the $\frac{1}{n}$-*periods*. Any half-integer characteristic is given by

$$\mathfrak{m} = \frac{1}{2}m = \frac{1}{2}\begin{pmatrix} m_1 & m_2 & \cdots & m_g \\ m_1' & m_2' & \cdots & m_g' \end{pmatrix}$$

where $m_i, m_i' \in \mathbb{Z}$. For $\gamma = \begin{bmatrix} \gamma' \\ \gamma'' \end{bmatrix} \in \frac{1}{2}\mathbb{Z}^{2g}/\mathbb{Z}^{2g}$ we define $e_*(\gamma) = (-1)^{4(\gamma')^t\gamma''}$. Then,

$$\theta[\gamma](-z,\tau) = e_*(\gamma)\theta[\gamma](z,\tau).$$

We say that γ is an **even** (resp. **odd**) characteristic if $e_*(\gamma) = 1$ (resp. $e_*(\gamma) = -1$). For any curve of genus g, there are $2^{g-1}(2^g+1)$ (respectively $2^{g-1}(2^g-1)$) even theta functions (respectively odd theta functions). Let $\mathfrak{a}$ be another half integer characteristic. We define $\mathfrak{m}\,\mathfrak{a}$ as follows.

$$\mathfrak{m}\,\mathfrak{a} = \frac{1}{2}\begin{pmatrix} t_1 & t_2 & \cdots & t_g \\ t_1' & t_2' & \cdots & t_g' \end{pmatrix}$$

where $t_i \equiv (m_i + a_i) \mod 2$ and $t_i' \equiv (m_i' + a_i') \mod 2$.

For the rest of this section we consider only characteristics $\frac{1}{2}q$ in which each of the elements q_i, q_i' is either 0 or 1. We use the following abbreviations

$$|\mathfrak{m}| = \sum_{i=1}^{g} m_i m_i', \qquad |\mathfrak{m},\mathfrak{a}| = \sum_{i=1}^{g}(m_i'a_i - m_i a_i'),$$

$$|\mathfrak{m},\mathfrak{a},\mathfrak{b}| = |\mathfrak{a},\mathfrak{b}| + |\mathfrak{b},\mathfrak{m}| + |\mathfrak{m},\mathfrak{a}|, \qquad \binom{\mathfrak{m}}{\mathfrak{a}} = e^{\pi i \sum_{j=1}^{g} m_j a_j'}.$$

The set of all half integer characteristics forms a group Γ which has 2^{2g} elements. We say that two half integer characteristics $\mathfrak{m}$ and $\mathfrak{a}$ are *syzygetic* (resp., *azygetic*) if $|\mathfrak{m},\mathfrak{a}| \equiv 0 \mod 2$ (resp., $|\mathfrak{m},\mathfrak{a}| \equiv 1 \mod 2$) and three half integer characteristics $\mathfrak{m}, \mathfrak{a}$, and $\mathfrak{b}$ are syzygetic if $|\mathfrak{m},\mathfrak{a},\mathfrak{b}| \equiv 0 \mod 2$.

A *Göpel group* G is a group of 2^r half integer characteristics where $r \leq g$ such that every two characteristics are syzygetic. The elements of the group G are formed by the sums of r fundamental characteristics; see [4, pg. 489] for details. Obviously, a Göpel group of order 2^r is isomorphic to C_2^r. The proof of the following lemma can be found on [4, pg. 490].

Lemma 1. *The number of different Göpel groups which have 2^r characteristics is*

$$\frac{(2^{2g}-1)(2^{2g-2}-1)\cdots(2^{2g-2r+2}-1)}{(2^r-1)(2^{r-1}-1)\cdots(2-1)}$$

If G is a Göpel group with 2^r elements, then it has 2^{2g-r} cosets. The cosets are called *Göpel systems* and denoted by $\mathfrak{a}G$, $\mathfrak{a} \in \Gamma$. Any three characteristics of a Göpel system are syzygetic. We can find a set of characteristics called a basis of the Göpel system which derives all its 2^r characteristics by taking only the combinations of any odd number of characteristics of the basis.

Lemma 2. *Let $g \geq 1$ be a fixed integer, r be as defined above and $\sigma = g - r$. Then there are $2^{\sigma-1}(2^\sigma+1)$ Göpel systems which consist of even characteristics only and there are $2^{\sigma-1}(2^\sigma-1)$ Göpel systems which consist of odd characteristics. The other $2^{2\sigma}(2^r-1)$ Göpel systems consist as many odd characteristics as even characteristics.*

Proof. The proof can be found on [4, pg. 492]. □

Corollary 3. *When $r = g$ we have only one (resp., 0) Göpel system which consists of even (resp., odd) characteristics.*

Proposition 4. *The following statements are true.*

$$\theta^2[\mathfrak{a}]\theta^2[\mathfrak{a}\mathfrak{h}] = \frac{1}{2^{g-1}} \sum_{\mathfrak{e}} e^{\pi i |\mathfrak{a}\mathfrak{e}|} \binom{\mathfrak{h}}{\mathfrak{a}\mathfrak{e}} \theta^2[\mathfrak{e}]\theta^2[\mathfrak{e}\mathfrak{h}] \tag{1}$$

$$\theta^4[\mathfrak{a}] + e^{\pi i |\mathfrak{a},\mathfrak{h}|}\theta^4[\mathfrak{a}\mathfrak{h}] = \frac{1}{2^{g-1}} \sum_{\mathfrak{e}} e^{\pi i |\mathfrak{a}\mathfrak{e}|} \{\theta^4[\mathfrak{e}] + e^{\pi i |\mathfrak{a},\mathfrak{h}|}\theta^4[\mathfrak{e}\mathfrak{h}]\} \tag{2}$$

where $\theta[e]$ is the theta constant corresponding to the characteristic e, $\mathfrak{a}$ and $\mathfrak{h}$ are any half integer characteristics and $\mathfrak{e}$ is an even characteristic such that $|\mathfrak{e}| \equiv |\mathfrak{e}\mathfrak{h}|$ mod 2. *There are $2 \cdot 2^{g-2}\,(2^{g-1}+1)$ such candidates for $\mathfrak{e}$.*

Proof. For the proof, see [4, pg. 524]. □

The statements given in the proposition above can be used to get identities among theta constants; see section 3.

2.1. **Cyclic curves with extra automorphisms.** A normal cyclic curve is an algebraic curve $\mathcal{X}$ such that there exist a normal cyclic subgroup $C_m \triangleleft \mathrm{Aut}(\mathcal{X})$ such that $g(\mathcal{X}/C_m) = 0$. Then $\bar{G} = G/C_m$ embeds as a finite subgroup of $PGL(2,\mathbb{C})$. An affine equation of a birational model of a cyclic curve can be given by the following

$$y^m = f(x) = \prod_{i=1}^{s} (x - \alpha_i)^{d_i},\ 0 < d_i < m. \tag{3}$$

Hyperelliptic curves are cyclic curves with $m = 2$. Note that when $0 < d_i$ for some i the curve is singular. A hyperelliptic curve $\mathcal{X}$ is a cover of order two of the projective line $\mathbb{P}^1$. Let z be the generator (the hyperelliptic involution) of the Galois group $Gal(\mathcal{X}/\mathbb{P}^1)$. It is known that $\langle z \rangle$ is a normal subgroup of the automorphism group $\mathrm{Aut}(\mathcal{X})$. Let $\mathcal{X} \longrightarrow \mathbb{P}^1$ be the degree 2 hyperelliptic projection. We can assume that infinity is a branch point. Let

$$B := \{\alpha_1, \alpha_2, \cdots, \alpha_{2g+1}\}$$

be the set of other branch points. Let $S = \{1, 2, \cdots, 2g+1\}$ be the index set of B and $\eta : S \longrightarrow \frac{1}{2}\mathbb{Z}^{2g}/\mathbb{Z}^{2g}$ be a map defined as follows;

$$\eta(2i-1) = \begin{bmatrix} 0 & \cdots & 0 & \frac{1}{2} & 0 & \cdots & 0 \\ \frac{1}{2} & \cdots & \frac{1}{2} & 0 & 0 & \cdots & 0 \end{bmatrix}$$

$$\eta(2i) = \begin{bmatrix} 0 & \cdots & 0 & \frac{1}{2} & 0 & \cdots & 0 \\ \frac{1}{2} & \cdots & \frac{1}{2} & \frac{1}{2} & 0 & \cdots & 0 \end{bmatrix}$$

where the nonzero element of the first row appears in i^{th} column. We define $\eta(\infty)$ to be $\begin{bmatrix} 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & 0 \end{bmatrix}$. For any $T \subset B$, we can define the half-integer characteristic as

$$\eta_T = \sum_{a_k \in T} \eta(k).$$

Let T^c denote the complement of T in B. Note that $\eta_B \in \mathbb{Z}^{2g}$. If we view η_T as an element of $\frac{1}{2}\mathbb{Z}^{2g}/\mathbb{Z}^{2g}$ then $\eta_T = \eta_{T^c}$. Let $\triangle$ denote the symmetric difference of

sets, that is $T \triangle R = (T \cup R) - (T \cap R)$. It can be shown that the set of subsets of B is a group under $\triangle$. We have the following group isomorphism

$$\{T \subset B \,|\, \#T \equiv g+1 \mod 2\}/T \cong \frac{1}{2}\mathbb{Z}^{2g}/\mathbb{Z}^{2g}.$$

For hyperelliptic curves, it is known that $2^{g-1}(2^g+1) - \binom{2g+1}{g}$ of the even theta constants are zero. The following theorem provides a condition on the characteristics in which theta characteristics become zero. The proof of the theorem can be found in [7, pg. 102].

Theorem 5. *Let $\mathcal{X}$ be a hyperelliptic curve, with a set B of branch points. Let S be the index set as above and U be the set of all odd values of S. Then for all $T \subset S$ with even cardinality, we have $\theta[\eta_T] = 0$ if and only if $\#(T \triangle U) \neq g+1$, where $\theta[\eta_T]$ is the theta constant corresponding to the characteristics η_T.*

Notice also that by parity, all odd theta constants are zero. There is a formula (so called Frobenius' theta formula) which half-integer theta characteristics for hyperelliptic curves satisfy.

Lemma 6 (Frobenius)**.** *For all $z_i \in \mathbb{C}^g$, $1 \leq i \leq 4$ such that $z_1 + z_2 + z_3 + z_4 = 0$ and for all $b_i \in \mathbb{Q}^{2g}$, $1 \leq i \leq 4$ such that $b_1 + b_2 + b_3 + b_4 = 0$, we have*

$$\sum_{j \in S \cup \{\infty\}} \epsilon_U(j) \prod_{i=1}^{4} \theta[b_i + \eta(j)](z_i) = 0,$$

where for any $A \subset B$,

$$\epsilon_A(k) = \begin{cases} 1 & \text{if } k \in A \\ -1 & \text{otherwise} \end{cases}$$

Proof. See [6, pg. 107]. □

A relationship between theta constants and the branch points of the hyperelliptic curve is given by Thomae's formula.

Lemma 7 (Thomae)**.** *For a non singular even half integer characteristics e corresponding to the partition of the branch points $\{1, 2, \cdots, 2(g+1)\} = \{i_1 < i_2 < \cdots < i_{g+1}\} \cup \{j_1 < j_2 < \cdots < j_{g+1}\}$, we have*

$$\theta[e](0;\tau)^8 = A \prod_{k<l} (\lambda_{i_k} - \lambda_{i_l})^2 (\lambda_{j_k} - \lambda_{j_l})^2.$$

See [6, pg. 128] for the description of A and [6, pg. 120] for the proof. Using Thomae's formula and Frobenius' theta identities we express the branch points of the hyperelliptic curves in terms of even theta constants.

3. Genus 2 curves

The automorphism group G of a genus 2 curve $\mathcal{X}$ in characteristic $\neq 2$ is isomorphic to $\mathbb{Z}_2$, $\mathbb{Z}_{10}$, V_4, D_8, D_{12}, $SL_2(3)$, $GL_2(3)$, or 2^+S_5. The case when $G \cong 2^+S_5$ occurs only in characteristic 5. If $G \cong SL_2(3)$ (resp., $GL_2(3)$) then $\mathcal{X}$ has equation $Y^2 = X^6 - 1$ (resp., $Y^2 = X(X^4 - 1)$). If $G \cong \mathbb{Z}_{10}$ then $\mathcal{X}$ has equation $Y^2 = X^6 - X$. For a fixed G from the list above, the locus of genus 2 curves with automorphism group G is an irreducible algebraic subvariety of $\mathcal{M}_2$. Such loci can be described in terms of the Igusa invariants.

For any genus 2 curve we have six odd theta characteristics and ten even theta characteristics. The following are the sixteen theta characteristics, where the first ten are even and the last six are odd. For simplicity, we denote them by $\theta_i = \begin{bmatrix} a \\ b \end{bmatrix}$ instead of $\theta_i \begin{bmatrix} a \\ b \end{bmatrix}(z,\tau)$ where $i = 1, \ldots, 10$ for the even theta functions.

$$\theta_1 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \theta_2 = \begin{bmatrix} 0 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \theta_3 = \begin{bmatrix} 0 & 0 \\ \frac{1}{2} & 0 \end{bmatrix}, \theta_4 = \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}, \theta_5 = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 0 \end{bmatrix},$$
$$\theta_6 = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}, \theta_7 = \begin{bmatrix} 0 & \frac{1}{2} \\ 0 & 0 \end{bmatrix}, \theta_8 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 0 \end{bmatrix}, \theta_9 = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}, \theta_{10} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix},$$

and the odd theta functions correspond to the following characteristics

$$\begin{bmatrix} 0 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{bmatrix}, \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} \end{bmatrix}$$

Consider the following Göpel group

$$G = \left\{ 0 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \mathfrak{m}_1 = \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}, \mathfrak{m}_2 = \begin{bmatrix} 0 & 0 \\ \frac{1}{2} & 0 \end{bmatrix}, \mathfrak{m}_1\mathfrak{m}_2 = \begin{bmatrix} 0 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \right\}.$$

Then, the corresponding Göpel systems are given by:

$$G = \left\{ \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ \frac{1}{2} & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \right\}$$
$$\mathfrak{b}_1 G = \left\{ \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \right\}$$
$$\mathfrak{b}_2 G = \left\{ \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}, \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \begin{bmatrix} 0 & \frac{1}{2} \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{bmatrix} \right\}$$
$$\mathfrak{b}_3 G = \left\{ \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} \end{bmatrix} \right\}$$

Notice that from all four cosets, only G has all even characteristics as noticed in Corollary 3. Using the Prop. 4 we have the following six identities for the above Göpel group.

$$\begin{cases} \theta_5^2\theta_6^2 &= \theta_1^2\theta_4^2 - \theta_2^2\theta_3^2 \\ \theta_5^4 + \theta_6^4 &= \theta_1^4 - \theta_2^4 - \theta_3^4 + \theta_4^4 \\ \theta_7^2\theta_9^2 &= \theta_1^2\theta_3^2 - \theta_2^2\theta_4^2 \\ \theta_7^4 + \theta_9^4 &= \theta_1^4 - \theta_2^4 + \theta_3^4 - \theta_4^4 \\ \theta_8^2\theta_{10}^2 &= \theta_1^2\theta_2^2 - \theta_3^2\theta_4^2 \\ \theta_8^4 + \theta_{10}^4 &= \theta_1^4 + \theta_2^4 - \theta_3^4 - \theta_4^4 \end{cases}$$

These identities express even theta constants in terms of four theta constants. We call them fundamental theta constants θ_1, θ_2, θ_3, θ_4.

Next we find the relation between theta characteristics and branch points of a genus two curve.

Lemma 8 (Picard)**.** *Let a genus 2 curve be given by*

$$Y^2 = X(X-1)(X-\lambda)(X-\mu)(X-\nu). \tag{4}$$

Then, λ, μ, ν can be written as follows:

$$\lambda = \frac{\theta_1^2\theta_3^2}{\theta_2^2\theta_4^2}, \quad \mu = \frac{\theta_3^2\theta_8^2}{\theta_4^2\theta_{10}^2}, \quad \nu = \frac{\theta_1^2\theta_8^2}{\theta_2^2\theta_{10}^2}. \tag{5}$$

Proof. There are several ways for relating λ, μ, ν to theta constants, depending on the ordering of the branch points of the curve. Let $B = \{\nu, \mu, \lambda, 1, 0\}$ be the branch points of the curves in this order and $U = \{\nu, \lambda, 0\}$ be the set of odd branch points. Using Lemma 7 we have the following set of equations of theta constants and branch points.

$$
\begin{array}{ll}
\theta_1^4 = A\,\nu\lambda(\mu-1)(\nu-\lambda) & \theta_2^4 = A\,\mu(\mu-1)(\nu-\lambda) \\
\theta_3^4 = A\,\mu\lambda(\mu-\lambda)(\nu-\lambda) & \theta_4^4 = A\,\nu(\nu-\lambda)(\mu-\lambda) \\
\theta_5^4 = A\,\lambda\mu(\nu-1)(\nu-\mu) & \theta_6^4 = A\,(\nu-\mu)(\nu-\lambda)(\mu-\lambda) \\
\theta_7^4 = A\,\mu(\nu-1)(\lambda-1)(\nu-\lambda) & \theta_8^4 = A\,\mu\nu(\nu-\mu)(\lambda-1) \\
\theta_9^4 = A\,\nu(\mu-1)(\lambda-1)(\mu-\lambda) & \theta_{10}^4 = A\,\lambda(\lambda-1)(\nu-\mu),
\end{array} \tag{6}
$$

where A is a constant. Choosing the appropriate equation from the set Eq. (6) we have the following:

$$
\lambda^2 = \left(\frac{\theta_1^2\theta_3^2}{\theta_2^2\theta_4^2}\right)^2 \quad \mu^2 = \left(\frac{\theta_3^2\theta_8^2}{\theta_4^2\theta_{10}^2}\right)^2 \quad \nu^2 = \left(\frac{\theta_1^2\theta_8^2}{\theta_2^2\theta_{10}^2}\right)^2.
$$

Each value for (λ, μ, ν) gives isomorphic genus 2 curves. Hence, we can choose

$$
\lambda = \frac{\theta_1^2\theta_3^2}{\theta_2^2\theta_4^2}, \quad \mu = \frac{\theta_3^2\theta_8^2}{\theta_4^2\theta_{10}^2}, \quad \nu = \frac{\theta_1^2\theta_8^2}{\theta_2^2\theta_{10}^2}.
$$

This completes the proof.

□

One of the main goals of this paper is to describe each locus of genus 2 curves with fixed automorphism group in terms of the fundamental theta constants. We have the following

Corollary 9. *Every genus two curve can be written in the form:*

$$
y^2 = x\,(x-1)\,\left(x - \frac{\theta_1^2\theta_3^2}{\theta_2^2\theta_4^2}\right)\,\left(x^2 - \frac{\theta_2^2\,\theta_3^2 + \theta_1^2\,\theta_4^2}{\theta_2^2\,\theta_4^2}\cdot\alpha\,x + \frac{\theta_1^2\theta_3^2}{\theta_2^2\theta_4^2}\,\alpha^2\right),
$$

where $\alpha = \frac{\theta_8^2}{\theta_{10}^2}$ *and in terms of* $\theta_1, \ldots, \theta_4$ *is given by*

$$
\alpha^2 + \frac{\theta_1^4 + \theta_2^4 - \theta_3^4 - \theta_4^4}{\theta_1^2\theta_2^2 - \theta_3^2\theta_4^2}\,\alpha + 1 = 0
$$

Furthermore, if $\alpha = \pm 1$ *then* $V_4 \hookrightarrow Aut(\mathcal{X})$.

Proof. Let's write the genus 2 curve in the following form:

$$
Y^2 = X(X-1)(X-\lambda)(X-\mu)(X-\nu)
$$

where λ, μ, ν are given by Eq. (5). Let $\alpha := \frac{\theta_8^2}{\theta_{10}^2}$. Then,

$$
\mu = \frac{\theta_3^2}{\theta_4^2}\,\alpha, \quad \nu = \frac{\theta_1^2}{\theta_2^2}\,\alpha
$$

Using the following two identities,

$$
\begin{aligned}
\theta_8^4 + \theta_{10}^4 &= \theta_1^4 + \theta_2^4 - \theta_3^4 - \theta_4^4 \\
\theta_8^2\theta_{10}^2 &= \theta_1^2\theta_2^2 - \theta_3^2\theta_4^2
\end{aligned} \tag{7}
$$

we have,

$$
\alpha^2 + \frac{\theta_1^4 + \theta_2^4 - \theta_3^4 - \theta_4^4}{\theta_1^2\theta_2^2 - \theta_3^2\theta_4^2}\,\alpha + 1 = 0 \tag{8}
$$

If $\alpha = \pm 1$ the $\mu\nu = \lambda$. It is well known that this implies that the genus 2 curve has an elliptic involution. Hence, $V_4 \hookrightarrow Aut(\mathcal{X})$. □

Remark 10. *i) From the above we have that $\theta_8^4 = \theta_{10}^4$ implies that $V_4 \hookrightarrow Aut(\mathcal{X})$. Lemma 15 determines a necessary and equivalent statement when $V_4 \hookrightarrow Aut(\mathcal{X})$.*

ii) The last part of the lemma above shows that if $\theta_8^4 = \theta_{10}^4$ then all coefficients of the genus 2 curve are given as rational functions of the 4 fundamental theta functions. Such fundamental theta functions determine the field of moduli of the given curve. Hence, the curve is defined over its field of moduli.

Corollary 11. *Let $\mathcal{X}$ be a genus 2 curve which has an elliptic involution. Then $\mathcal{X}$ is defined over its field of moduli.*

This was the main result of [1].

3.1. **Describing the locus of genus two curves with fixed automorphism group by theta constants.** The locus $\mathcal{L}_2$ of genus 2 curves $\mathcal{X}$ which have an elliptic involution is a closed subvariety of $\mathcal{M}_2$. Let $W = \{\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2\}$ be the set of roots of the binary sextic and A and B be subsets of W such that $W = A \cup B$ and $|A \cap B| = 2$. We define the cross ratio of the two pairs $z_1, z_2; z_3, z_4$ by

$$(z_1, z_2; z_3, z_4) = \frac{z_1; z_3, z_4}{z_2; z_3, z_4} = \frac{z_1 - z_3}{z_1 - z_4} : \frac{z_2 - z_3}{z_2 - z_4}.$$

Take $A = \{\alpha_1, \alpha_2, \beta_1, \beta_2\}$ and $B = \{\gamma_1, \gamma_2, \beta_1, \beta_2\}$. Jacobi [2] gives a description of $\mathcal{L}_2$ in terms of the cross ratios of the elements of W.

$$\frac{\alpha_1 - \beta_1}{\alpha_1 - \beta_2} : \frac{\alpha_2 - \beta_1}{\alpha_2 - \beta_2} = \frac{\gamma_1 - \beta_1}{\gamma_1 - \beta_2} : \frac{\gamma_2 - \beta_1}{\gamma_2 - \beta_2}$$

We recall that the following identities hold for cross ratios:

$$(\alpha_1, \alpha_2\,; \beta_1, \beta_2) = (\alpha_2, \alpha_1; \beta_2, \beta_1) = (\beta_1, \beta_2; \alpha_1, \alpha_2) = (\beta_2, \beta_1; \alpha_2, \alpha_1)$$

and

$$(\alpha_1, \alpha_2; \infty, \beta_2) = (\infty, \beta_2; \alpha_1, \alpha_2) = (\beta_2; \alpha_2, \alpha_1)$$

Next we want to use this result to determine relations among theta functions for a genus 2 curve in the locus $\mathcal{L}_2$. Let $\mathcal{X}$ be any genus 2 curve given by equation

$$Y^2 = X(X-1)(X-a_1)(X-a_2)(X-a_3)$$

We take $\infty \in A \cap B$. Then there are five cases for $\alpha \in A \cap B$, where α is an element of the set$\{0, 1, a_1, a_2, a_3\}$. For each of these cases there are three possible relationships for cross ratios as described below:

i) $A \cap B = \{0, \infty\}$: The possible cross ratios are

$$(a_1, 1; \infty, 0) = (a_3, a_2; \infty, 0)$$

$$(a_2, 1; \infty, 0) = (a_1, a_3; \infty, 0)$$

$$(a_1, 1; \infty, 0) = (a_2, a_3; \infty, 0)$$

ii) $A \cap B = \{1, \infty\}$: The possible cross ratios are

$$(a_1, 0; \infty, 1) = (a_2, a_3; \infty, 1)$$

$$(a_1, 0; \infty, 1) = (a_3, a_2; \infty, 1)$$

$$(a_2, 0; \infty, 1) = (a_1, a_3; \infty, 1)$$

iii) $A \cap B = \{a_1, \infty\}$: The possible cross ratios are

$$(1, 0; \infty, a_1) = (a_3, a_2; \infty, a_1)$$
$$(a_2, 0; \infty, a_1) = (1, a_3; \infty, a_1)$$
$$(1, 0; \infty, a_1) = (a_2, a_3; \infty, a_1)$$

iv) $A \cap B = \{a_2, \infty\}$: The possible cross ratios are

$$(1, 0; \infty, a_2) = (a_1, a_3; \infty, a_2)$$
$$(1, 0; \infty, a_2) = (a_3, a_1; \infty, a_2)$$
$$(a_1, 0; \infty, a_2) = (1, a_3; \infty, a_2)$$

v) $A \cap B = \{a_3, \infty\}$: The possible cross ratios are

$$(a_1, 0; \infty, a_3) = (1, a_2; \infty, a_3)$$
$$(1, 0; \infty, a_3) = (a_2, a_1; \infty, a_3)$$
$$(1, 0; \infty, a_3) = (a_1, a_2; \infty, a_3)$$

We summarize these relationships in the following table:

	Cross ratio	$f(a_1, a_2, a_3) = 0$	theta constants
1	$(1, 0; \infty, a_1) = (a_3, a_2; \infty, a_1)$	$a_1a_2 + a_1 - a_3a_1 - a_2$	$-\theta_1^2\theta_3^2\theta_8^2\theta_2^2 - \theta_1^2\theta_2^2\theta_4^2\theta_{10}^2 + \theta_1^4\theta_3^2\theta_{10}^2 + \theta_3^2\theta_2^4\theta_{10}^2$
2	$(a_2, 0; \infty, a_1) = (1, a_3; \infty, a_1)$	$a_1a_2 - a_1 + a_3a_1 - a_3a_2$	$\theta_3^2\theta_8^2\theta_2^2\theta_4^2 - \theta_2^2\theta_4^4\theta_{10}^2 + \theta_1^2\theta_3^2\theta_4^2\theta_{10}^2 - \theta_3^4\theta_2^2\theta_{10}^2$
3	$(1, 0; \infty, a_1) = (a_2, a_3; \infty, a_1)$	$a_1a_2 - a_1 - a_3a_1 + a_3$	$-\theta_8^4\theta_2^2\theta_2^2 + \theta_8^2\theta_2^2\theta_{10}^2\theta_4^2 + \theta_1^2\theta_3^2\theta_8^2\theta_{10}^2 - \theta_3^2\theta_2^2\theta_{10}^4$
4	$(1, 0; \infty, a_2) = (a_1, a_3; \infty, a_2)$	$a_1a_2 - a_2 - a_3a_2 + a_3$	$-\theta_1^2\theta_8^4\theta_4^2 - \theta_1^2\theta_{10}^4\theta_4^2 + \theta_8^2\theta_2^2\theta_{10}^2\theta_4^2 + \theta_1^2\theta_3^2\theta_8^2\theta_{10}^2$
5	$(1, 0; \infty, a_2) = (a_3, a_1; \infty, a_2)$	$a_1a_2 - a_1 + a_2 - a_3a_2$	$-\theta_1^2\theta_8^2\theta_3^2\theta_4^2 + \theta_1^2\theta_{10}^2\theta_4^4 + \theta_1^2\theta_3^2\theta_{10}^2 - \theta_3^2\theta_2^2\theta_{10}^2\theta_4^2$
6	$(a_1, 0; \infty, a_2) = (1, a_3; \infty, a_2)$	$a_1a_2 - a_3a_1 - a_2 + a_3a_2$	$-\theta_1^2\theta_8^2\theta_2^2\theta_4^2 + \theta_1^4\theta_{10}^2\theta_4^2 - \theta_1^2\theta_3^2\theta_2^2\theta_{10}^2 + \theta_2^4\theta_4^2\theta_{10}^2$
7	$(a_1, 0; \infty, a_3) = (1, a_2; \infty, a_3)$	$a_1a_2 - a_3a_1 - a_3a_2 + a_3$	$-\theta_8^4\theta_2^2\theta_4^2 + \theta_1^2\theta_8^2\theta_{10}^2\theta_4^2 - \theta_2^2\theta_{10}^4\theta_4^2 + \theta_3^2\theta_8^2\theta_2^2\theta_{10}^2$
8	$(1, 0; \infty, a_3) = (a_2, a_1; \infty, a_3)$	$a_3a_1 - a_1 - a_3a_2 + a_3$	$\theta_8^4 - \theta_{10}^4$
9	$(1, 0; \infty, a_3) = (a_1, a_2; \infty, a_3)$	$a_3a_1 + a_2 - a_3 - a_3a_2$	$\theta_1^4\theta_8^2\theta_4^2 - \theta_1^2\theta_2^2\theta_4^2\theta_{10}^2 - \theta_1^2\theta_3^2\theta_8^2\theta_2^2 + \theta_8^2\theta_2^4\theta_4^2$
10	$(a_1, 0; \infty, 1) = (a_2, a_3; \infty, 1)$	$-a_1 + a_3a_1 + a_2 - a_3$	$\theta_3^4\theta_2^2\theta_8^2 - \theta_1^2\theta_8^2\theta_2^2\theta_4^2 - \theta_1^2\theta_3^2\theta_2^2\theta_{10}^2 + \theta_3^2\theta_8^2\theta_4^4$
11	$(a_1, 0; \infty, 1) = (a_3, a_2; \infty, 1)$	$a_1a_2 - a_1 - a_2 + a_3$	$\theta_1^2\theta_8^4\theta_3^2 - \theta_1^2\theta_8^2\theta_{10}^2\theta_4^2 + \theta_1^2\theta_3^2\theta_{10}^4 - \theta_3^2\theta_8^2\theta_2^2\theta_{10}^2$
12	$(a_2, 0; \infty, 1) = (a_1, a_3; \infty, 1)$	$a_1 - a_2 + a_3a_2 - a_3$	$\theta_1^2\theta_8^2\theta_4^4 - \theta_1^2\theta_3^2\theta_4^2\theta_{10}^2 + \theta_1^2\theta_3^4\theta_8^2 - \theta_3^2\theta_8^2\theta_2^2\theta_4^2$
13	$(a_1, 1; \infty, 0) = (a_3, a_2; \infty, 0)$	$a_1a_2 - a_3$	$\theta_8^4 - \theta_{10}^4$
14	$(a_2, 1; \infty, 0) = (a_1, a_3; \infty, 0)$	$a_1 - a_3a_2$	$\theta_3^4 - \theta_4^4$
15	$(a_1, 1; \infty, 0) = (a_2, a_3; \infty, 0)$	$a_3a_1 - a_2$	$\theta_1^4 - \theta_2^4$

TABLE 1. Relation of theta functions and cross ratios

Lemma 12. *Let $\mathcal{X}$ be a genus 2 curve. Then $Aut(\mathcal{X}) \cong V_4$ if and only if the theta functions of $\mathcal{X}$ satisfy*

$$\begin{aligned}
&(\theta_1^4-\theta_2^4)(\theta_3^4-\theta_4^4)(\theta_8^4-\theta_{10}^4)(-\theta_1^2\theta_3^2\theta_8^2\theta_2^2-\theta_1^2\theta_2^2\theta_4^2\theta_{10}^2+\theta_1^4\theta_3^2\theta_{10}^2+\theta_3^2\theta_2^4\theta_{10}^2)\\
&(\theta_3^2\theta_8^2\theta_2^2\theta_4^2-\theta_2^2\theta_4^4\theta_{10}^2+\theta_1^2\theta_3^2\theta_4^2\theta_{10}^2-\theta_3^4\theta_2^2\theta_{10}^2)(-\theta_8^4\theta_3^2\theta_2^2+\theta_8^2\theta_2^2\theta_{10}^2\theta_4^2+\theta_1^2\theta_3^2\theta_8^2\theta_{10}^2-\theta_3^2\theta_2^2\theta_{10}^4)\\
&(-\theta_1^2\theta_8^4\theta_4^2-\theta_1^2\theta_{10}^4\theta_4^2+\theta_8^2\theta_2^2\theta_{10}^2\theta_4^2+\theta_1^2\theta_3^2\theta_8^2\theta_{10}^2)(-\theta_1^2\theta_8^2\theta_3^2\theta_4^2+\theta_1^2\theta_{10}^2\theta_4^4+\theta_1^2\theta_3^4\theta_{10}^2-\theta_3^2\theta_2^2\theta_{10}^2\theta_4^2)\\
&(-\theta_1^2\theta_8^2\theta_2^2\theta_4^2+\theta_1^4\theta_{10}^2\theta_4^2-\theta_1^2\theta_3^2\theta_2^2\theta_{10}^2+\theta_2^2\theta_4^4\theta_{10}^2)(-\theta_8^4\theta_2^2\theta_4^2+\theta_1^2\theta_8^2\theta_{10}^2\theta_4^2-\theta_2^2\theta_{10}^4\theta_4^2+\theta_3^2\theta_8^2\theta_2^2\theta_{10}^2)\\
&(\theta_1^4\theta_8^2\theta_4^2-\theta_1^2\theta_2^2\theta_4^2\theta_{10}^2-\theta_1^2\theta_3^2\theta_8^2\theta_2^2+\theta_8^2\theta_4^4\theta_2^2)(\theta_1^4\theta_3^2\theta_8^2-\theta_1^2\theta_8^2\theta_2^2\theta_4^2-\theta_1^2\theta_3^2\theta_2^2\theta_{10}^2+\theta_3^2\theta_8^2\theta_2^4)\\
&(\theta_1^2\theta_8^4\theta_3^2-\theta_1^2\theta_8^2\theta_{10}^2\theta_4^2+\theta_1^2\theta_3^2\theta_{10}^4-\theta_3^2\theta_8^2\theta_2^2\theta_{10}^2)(\theta_1^2\theta_8^2\theta_4^4-\theta_1^2\theta_3^2\theta_4^2\theta_{10}^2+\theta_1^2\theta_3^4\theta_8^2-\theta_3^2\theta_8^2\theta_2^2\theta_4^2)=0
\end{aligned} \tag{9}$$

However, we are unable to get a similar result for cases D_8 or D_{12} by this argument. Instead, we will use the invariants of genus 2 curves and a more computational approach. In the process, we will offer a different proof of the lemma above.

Lemma 13. *i) The locus $\mathcal{L}_2$ of genus 2 curves $\mathcal{X}$ which have a degree 2 elliptic subcover is a closed subvariety of $\mathcal{M}_2$. The equation of $\mathcal{L}_2$ is given by*

$$\begin{aligned}
&8748J_{10}J_2^4J_6^2-507384000J_{10}^2J_4^2J_2-19245600J_{10}^2J_4J_2^3-592272J_{10}J_4^4J_2^2+77436J_{10}J_4^3J_2^4\\
&-81J_2^3J_6^4-3499200J_{10}J_2J_6^3+4743360J_{10}J_4^3J_2J_6-870912J_{10}J_4^2J_2^3J_6+3090960J_{10}J_4J_2^2J_6^2\\
&-78J_2^5J_4^5-125971200000J_{10}^3+384J_4^6J_6+41472J_{10}J_4^5+159J_4^6J_2^3-236196J_{10}^2J_2^5-80J_4^7J_2\\
&-47952J_2J_4J_6^4+104976000J_{10}^2J_2^2J_6-1728J_4^5J_2^2J_6+6048J_4^4J_2J_6^2-9331200J_{10}J_4^2J_6^2\\
&+12J_2^6J_4^3J_6+29376J_2^2J_4^2J_6^3-8910J_2^3J_4^3J_6^2-2099520000J_{10}^2J_4J_6+31104J_6^5-6912J_4^3J_6^34\\
&-J_2^7J_4^4-5832J_{10}J_2^5J_4J_6-54J_2^5J_4^2J_6^2+108J_2^4J_4J_6^3+972J_{10}J_2^6J_4^2+1332J_2^4J_4^4J_6=0
\end{aligned} \tag{10}$$

ii) The locus of genus 2 curves $\mathcal{X}$ with $Aut(\mathcal{X}) \cong D_8$ is given by the equation of $\mathcal{L}_2$ and

$$1706J_4^2J_2^2+2560J_4^3+27J_4J_2^4-81J_2^3J_6-14880J_2J_4J_6+28800J_6^2=0 \tag{11}$$

iii) The locus of genus 2 curves $\mathcal{X}$ with $Aut(\mathcal{X}) \cong D_{12}$ is

$$\begin{aligned}
&-J_4J_2^4+12J_2^3J_6-52J_4^2J_2^2+80J_4^3+960J_2J_4J_6-3600J_6^2=0\\
&864J_{10}J_2^5+3456000J_{10}J_4^2J_2-43200J_{10}J_4J_2^3-2332800000J_{10}^2-J_4^2J_2^6\\
&-768J_4^4J_2^2+48J_4^3J_2^4+4096J_4^5=0
\end{aligned} \tag{12}$$

Our goal is to express each of the above loci in terms of the theta characteristics. We obtain the following result.

Theorem 14. *Let $\mathcal{X}$ be a genus 2 curve. Then the following hold:*
i) $Aut(\mathcal{X}) \cong V_4$ if and only if the relations of theta functions given Eq. (9) *holds.*
ii) $Aut(\mathcal{X}) \cong D_8$ if and only if Eq. (1) in [10] *is satisfied.*
iii) $Aut(\mathcal{X}) \cong D_{12}$ if and only if Eq. (2) in [10] *is satisfied.*

Proof. Part i) of the theorem is Lemma 12. Here we give a somewhat different proof. Assume that $\mathcal{X}$ is a genus 2 curve with equation

$$Y^2 = X(X-1)(X-a_1)(X-a_2)(X-a_3)$$

whose classical invariants satisfy Eq. (10). Expressing the classical invariants of $\mathcal{X}$ in terms of a_1, a_2, a_3, substituting them into (10), and factoring the resulting

equation yields

$$
\begin{aligned}
&(a_1a_2-a_2-a_3a_2+a_3)^2(a_1a_2-a_1+a_3a_1-a_3a_2)^2(a_1a_2-a_3a_1-a_3a_2+a_3)^2\\
&(a_3a_1-a_1-a_3a_2+a_3)^2(a_1a_2+a_1-a_3a_1-a_2)^2(a_1a_2-a_1-a_3a_1+a_3)^2\\
&(a_3a_1+a_2-a_3-a_3a_2)^2(-a_1+a_3a_1+a_2-a_3)^2(a_1a_2-a_1-a_2+a_3)^2\\
&(a_1a_2-a_1+a_2-a_3a_2)^2(a_1-a_2+a_3a_2-a_3)^2(a_1a_2-a_3a_1-a_2+a_3a_2)^2\\
&(a_1a_2-a_3)^2(a_1-a_3a_2)^2(a_3a_1-a_2)^2=0
\end{aligned}
\tag{13}
$$

It is no surprise that we get the 15 factors of Table 1. The relations of theta constants follow from the table. ii) Let $\mathcal{X}$ be a genus 2 curve which has an elliptic involution. Then $\mathcal{X}$ is isomorphic to a curve with equation

$$Y^2=X(X-1)(X-a_1)(X-a_2)(X-a_1a_2).$$

If $\mathrm{Aut}(\mathcal{X})\cong D_8$ then the $SL_2(k)$-invariants of such curve must satisfy Eq. (11). Then, we get the equation in terms of a_1,a_2. By writing the relation $a_3=a_1a_2$ in terms of theta constants, we get $\theta_4^4=\theta_3^4$. All the results above lead to part ii) of the theorem. iii) The proof of this part is similar to part ii). □

We would like to express the conditions of the previous lemma in terms of the fundamental theta constants only.

Lemma 15. *Let $\mathcal{X}$ be a genus 2 curve. Then we have the following:*

i): *$V_4\hookrightarrow Aut(\mathcal{X})$ if and only if the fundamental theta constants of $\mathcal{X}$ satisfy*

$$
\begin{aligned}
&\left(\theta_3^4-\theta_4^4\right)\left(\theta_1^4-\theta_3^4\right)\left(\theta_2^4-\theta_4^4\right)\left(\theta_1^4-\theta_4^4\right)\left(\theta_3^4-\theta_2^4\right)\left(\theta_1^4-\theta_2^4\right)\\
&\left(-\theta_4^2+\theta_3^2+\theta_1^2-\theta_2^2\right)\left(\theta_4^2-\theta_3^2+\theta_1^2-\theta_2^2\right)\left(-\theta_4^2-\theta_3^2+\theta_2^2+\theta_1^2\right)\left(\theta_4^2+\theta_3^2+\theta_2^2+\theta_1^2\right)\\
&\left({\theta_1}^4{\theta_2}^4+{\theta_3}^4{\theta_2}^4+{\theta_1}^4{\theta_3}^4-2\,\theta_1^2\theta_2^2\theta_3^2\theta_4^2\right)\left(-{\theta_3}^4{\theta_2}^4-{\theta_2}^4{\theta_4}^4-{\theta_3}^4{\theta_4}^4+2\,\theta_1^2\theta_2^2\theta_3^2\theta_4^2\right)\\
&\left({\theta_2}^4{\theta_4}^4+{\theta_1}^4{\theta_2}^4+{\theta_1}^4{\theta_4}^4-2\,\theta_1^2\theta_2^2\theta_3^2\theta_4^2\right)\left({\theta_1}^4{\theta_4}^4+{\theta_3}^4{\theta_4}^4+{\theta_1}^4{\theta_3}^4-2\,\theta_1^2\theta_2^2\theta_3^2\theta_4^2\right)=0
\end{aligned}
\tag{14}
$$

ii: *$D_8\hookrightarrow Aut(\mathcal{X})$ if and only if the fundamental theta constants of $\mathcal{X}$ satisfy Eq. (3) in* [10]

iii: *$D_6\hookrightarrow Aut(\mathcal{X})$ if and only if the fundamental theta constants of $\mathcal{X}$ satisfy Eq. (4) in* [10]

Proof. Notice that Eq. (9) contains only $\theta_1,\theta_2,\theta_3,\theta_4,\theta_8$ and θ_{10}. Using Eq. (7), we can eliminate θ_8 and θ_{10} from Eq. (9). The J_{10} invariant of any genus two curve is given by the following in terms of theta constants:

$$J_{10}=\frac{\theta_1^{12}\theta_3^{12}}{\theta_2^{28}\theta_4^{28}\theta_{10}^{40}}\,(\theta_1^2\theta_2^2-\theta_3^2\theta_4^2)^{12}(\theta_1^2\theta_4^2-\theta_2^2\theta_3^2)^{12}(\theta_1^2\theta_3^2-\theta_2^2\theta_4^2)^{12}.$$

Since $J_{10}\neq 0$ we can cancel the factors $(\theta_1^2\theta_2^2-\theta_3^2\theta_4^2)$, $(\theta_1^2\theta_4^2-\theta_2^2\theta_3^2)$ and $(\theta_1^2\theta_3^2-\theta_2^2\theta_4^2)$ from the equation of V_4 locus. The result follows from Theorem 14. The proof of part ii) and iii) is similar and we avoid details. □

Remark 16. *i) For the other two loci, we can also obtain equations in terms of the fundamental theta constants. However, such equations are big and we don't display them here.*

ii) By using Frobenius's relations we get

$$J_{10}=\frac{(\theta_1\theta_3)^{12}}{(\theta_2\theta_4)^{28}\,\theta_{10}^{16}}\,(\theta_5\theta_6\theta_7\theta_8\theta_9)^{24}$$

Hence, $\theta_i \neq 0$ *for* $i = 1, 3, 5, \ldots 9$.

4. Genus 3 cyclic curves

For genus 3 we have hyperelliptic and non-hyperelliptic algebraic curves. The following table gives all possible genus 3 cyclic algebraic curves; see [5] for details. The first 11 cases are for the hyperelliptic curves and the last 12 cases are for the non-hyperelliptic curves.

	$\mathrm{Aut}(\mathcal{X}_g)$	equation	Id.
1	$\mathbb{Z}_2$	$y^2 = x(x-1)(x^5 + ax^4 + bx^3 + cx^2 + dx + e)$	$(2,1)$
2	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$y^2 = x^8 + a_3x^6 + a_2x^4 + a_1x^2 + 1$	$(4,2)$
3	$\mathbb{Z}_4$	$y^2 = x(x^2-1)(x^4 + ax^2 + b)$	$(4,1)$
4	$\mathbb{Z}_{14}$	$y^2 = x^7 - 1$	$(14,2)$
5	$\mathbb{Z}_2^3$	$y^2 = (x^4 + ax^2 + 1)(x^4 + bx^2 + 1)$	$(8,5)$
6	$\mathbb{Z}_2 \times D_8$	$y^2 = x^8 + ax^4 + 1$	$(16,11)$
7	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$y^2 = (x^4 - 1)(x^4 + ax^2 + 1)$	$(8,2)$
8	D_{12}	$y^2 = x(x^6 + ax^3 + 1)$	$(12,4)$
9	U_6	$y^2 = x(x^6 - 1)$	$(24,5)$
10	V_8	$y^2 = x^8 - 1$	$(32,9)$
11	$\mathbb{Z}_2 \times S_4$	$y^2 = x^8 + 14x^2 + 1$	$(48,48)$
12	V_4	$x^4 + y^4 + ax^2y^2 + bx^2 + cy^2 + 1 = 0$	(4,2)
13	D_8	take $b = c$	(8,3)
14	S_4	take $a = b = c$	(24,12)
15	$C_4^2 \rtimes S_3$	take $a = b = c = 0$ or $y^4 = x(x^2 - 1)$	(96,64)
16	16	$y^4 = x(x-1)(x-t)$	(16,13)
17	48	$y^4 = x^3 - 1$	(48,33)
18	C_3	$y^3 = x(x-1)(x-s)(x-t)$	(3,1)
19	C_6	take $s = 1 - t$	(6,2)
20	C_9	$y^3 = x(x^3 - 1)$	(9,1)
21	$L_3(2)$	$x^3y + y^3z + z^3x = 0$	(168,42)
22	S_3	$a(x^4 + y^4 + z^4) + b(x^2y^2 + x^2z^2 + y^2z^2) + c(x^2yz + y^2xz + z^2xy) = 0$	(6,1)
23	C_2	$x^4 + x^2(y^2 + az^2) + by^4 + cy^3z + dy^2z^2 + eyz^3 + gz^4 = 0$, either $e = 1$ or $g = 1$	(2,1)

Table 2. The list of automorphism groups of genus 3 and their equations

4.1. **Theta functions for hyperelliptic curves.** For genus three hyperelliptic curve we have 28 odd theta characteristics and 36 even theta characteristics. The following shows the corresponding characteristics for each theta function. The first 36 are for the even functions and the last 28 are for the odd functions. For simplicity, we denote them by $\theta_i \begin{bmatrix} a \\ b \end{bmatrix}$ instead of $\theta_i \begin{bmatrix} a \\ b \end{bmatrix}(z,\tau)$.

$$\theta_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},\ \theta_2 = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix},\ \theta_3 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix},\ \theta_4 = \begin{bmatrix} 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 \end{bmatrix},$$
$$\theta_5 = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix},\ \theta_6 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix},\ \theta_7 = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 \end{bmatrix},\ \theta_8 = \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{bmatrix},$$
$$\theta_9 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix},\ \theta_{10} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},\ \theta_{11} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix},\ \theta_{12} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$$
$$\theta_{13} = \begin{bmatrix} 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix},\ \theta_{14} = \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix},\ \theta_{15} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix},\ \theta_{16} = \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix},$$
$$\theta_{17} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix},\ \theta_{18} = \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix},\ \theta_{19} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix},\ \theta_{20} = \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix},$$
$$\theta_{21} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix},\ \theta_{22} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix},\ \theta_{23} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix},\ \theta_{24} = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$$
$$\theta_{25} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix},\ \theta_{26} = \begin{bmatrix} 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix},\ \theta_{27} = \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \end{bmatrix},\ \theta_{28} = \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix},$$
$$\theta_{29} = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix},\ \theta_{30} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix},\ \theta_{31} = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{bmatrix},\ \theta_{32} = \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 \end{bmatrix},$$
$$\theta_{33} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix},\ \theta_{34} = \begin{bmatrix} 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix},\ \theta_{35} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix},\ \theta_{36} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
$$\theta_{37} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 \end{bmatrix},\ \theta_{38} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix},\ \theta_{39} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} \end{bmatrix},\ \theta_{40} = \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix},$$
$$\theta_{41} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix},\ \theta_{42} = \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix},\ \theta_{43} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 \end{bmatrix},\ \theta_{44} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{bmatrix},$$
$$\theta_{45} = \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} \end{bmatrix},\ \theta_{46} = \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix},\ \theta_{47} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix},\ \theta_{48} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$$
$$\theta_{49} = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix},\ \theta_{50} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix},\ \theta_{51} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix},\ \theta_{52} = \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix},$$
$$\theta_{53} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} \end{bmatrix},\ \theta_{54} = \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix},\ \theta_{55} = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} \end{bmatrix},\ \theta_{56} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix},$$
$$\theta_{57} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \end{bmatrix},\ \theta_{58} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{bmatrix},\ \theta_{59} = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 \end{bmatrix},\ \theta_{60} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$
$$\theta_{61} = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix},\ \theta_{62} = \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix},\ \theta_{63} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix},\ \theta_{64} = \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

It can be shown that one of the corresponding even theta constants is zero. Let's pick $S = \{1,2,3,4,5,6,7\}$ and $U = \{1,3,5,7\}$. Let $T = U$. Then, by Theorem 5 the theta constant corresponding to the characteristic $\eta_T = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$ is zero. That is $\theta_{12}(0) = 0$. Next, we give the relation between theta characteristics and branch points of the genus three hyperelliptic curve. Let $B = \{a_1, a_2, a_3, a_4, a_5, 1, 0\}$ be the finite branch points of the curves and $U = \{a_1, a_3, a_5, 0\}$ be the set of odd branch points.

Lemma 17. *Any genus 3 hyperelliptic curve is isomorphic to a curve given by the equation*

$$Y^2 = X(X-1)(X-a_1)(X-a_2)(X-a_3)(X-a_4)(X-a_5),$$

where

$$a_1 = \frac{\theta_{31}^2\theta_{21}^2}{\theta_{34}^2\theta_{24}^2},\ a_2 = \frac{\theta_{31}^2\theta_{13}^2}{\theta_9^2\theta_{24}^2},\ a_3 = \frac{\theta_{11}^2\theta_{31}^2}{\theta_{24}^2\theta_6^2},\ a_4 = \frac{\theta_{21}^2\theta_7^2}{\theta_{15}^2\theta_{34}^2},\ a_5 = \frac{\theta_{13}^2\theta_1^2}{\theta_{26}^2\theta_9^2}.$$

Proof. By using Lemma 7 we have the following set of equation of theta constants and branch points which are ordered $a_1, a_2, a_3, a_4, a_5, 0, 1, \infty$. We use the notation (i, j) for $(a_i - a_j)$.

$$\begin{aligned}
{\theta_1}^4 &= A\ (1,6)\,(3,6)\,(5,6)\,(1,3)\,(1,5)\,(3,5)\,(2,4)\,(2,7)\,(4,7)\\
{\theta_2}^4 &= A\ (3,6)\,(5,6)\,(3,5)\,(1,2)\,(1,4)\,(2,4)\,(3,7)\,(5,7)\\
{\theta_3}^4 &= A\ (3,6)\,(4,6)\,(3,4)\,(1,2)\,(1,5)\,(2,5)\,(1,7)\,(2,7)\,(5,7)\\
{\theta_4}^4 &= A\ (2,6)\,(3,6)\,(5,6)\,(2,3)\,(2,5)\,(3,5)\,(1,4)\,(1,7)\,(4,7)\\
{\theta_5}^4 &= A\ (4,6)\,(5,6)\,(4,5)\,(1,2)\,(1,3)\,(2,3)\,(1,7)\,(2,7)\,(3,7)\\
{\theta_6}^4 &= A\ (1,6)\,(2,6)\,(3,4)\,(3,5)\,(4,5)\,(1,2)\,(1,7)\,(2,7)\\
{\theta_7}^4 &= A\ (2,6)\,(3,6)\,(4,6)\,(1,5)\,(2,3)\,(2,4)\,(3,4)\,(1,7)\,(5,7)\\
{\theta_8}^4 &= A\ (2,6)\,(3,6)\,(2,3)\,(1,4)\,(1,5)\,(4,5)\,(1,7)\,(4,7)\,(5,7)\\
{\theta_9}^4 &= A\ (1,6)\,(3,6)\,(1,3)\,(2,4)\,(2,5)\,(4,5)\,(1,7)\,(3,7)\\
{\theta_{10}}^4 &= A\ (3,6)\,(5,6)\,(3,5)\,(1,2)\,(1,4)\,(2,4)\,(1,7)\,(2,7)\,(4,7)\\
{\theta_{11}}^4 &= A\ (3,6)\,(4,6)\,(5,6)\,(3,4)\,(3,5)\,(4,5)\,(1,2)\,(1,7)\,(2,7)\\
{\theta_{13}}^4 &= A\ (2,6)\,(4,6)\,(5,6)\,(1,3)\,(2,4)\,(2,5)\,(4,5)\,(1,7)\,(3,7)\\
{\theta_{14}}^4 &= A\ (2,6)\,(5,6)\,(2,5)\,(1,3)\,(1,4)\,(3,4)\,(1,7)\,(3,7)\,(4,7)\\
{\theta_{15}}^4 &= A\ (1,6)\,(5,6)\,(1,5)\,(2,3)\,(2,4)\,(3,4)\,(1,7)\,(5,7)\\
{\theta_{16}}^4 &= A\ (1,6)\,(2,3)\,(2,4)\,(2,5)\,(3,4)\,(3,5)\,(4,5)\,(1,7)\\
{\theta_{17}}^4 &= A\ (1,6)\,(4,6)\,(2,3)\,(2,5)\,(3,5)\,(1,4)\,(1,7)\,(4,7)\\
{\theta_{18}}^4 &= A\ (2,6)\,(4,6)\,(1,3)\,(1,5)\,(3,5)\,(2,4)\,(1,7)\,(3,7)\,(5,7)\\
{\theta_{19}}^4 &= A\ (3,6)\,(4,6)\,(1,2)\,(1,5)\,(2,5)\,(3,4)\,(3,7)\,(4,7)\\
{\theta_{20}}^4 &= A\ (2,6)\,(1,3)\,(1,4)\,(1,5)\,(3,4)\,(3,5)\,(4,5)\,(2,7)\\
{\theta_{21}}^4 &= A\ (1,6)\,(4,6)\,(5,6)\,(1,4)\,(1,5)\,(4,5)\,(2,3)\,(2,7)\,(3,7)\\
{\theta_{22}}^4 &= A\ (1,6)\,(3,6)\,(4,6)\,(1,3)\,(1,4)\,(3,4)\,(2,5)\,(2,7)\,(5,7)\\
{\theta_{23}}^4 &= A\ (1,6)\,(2,6)\,(3,4)\,(3,5)\,(4,5)\,(1,2)\,(3,7)\,(4,7)\,(5,7)\\
{\theta_{24}}^4 &= A\ (4,6)\,(5,6)\,(1,2)\,(1,3)\,(2,3)\,(4,5)\,(4,7)\,(5,7)\\
{\theta_{25}}^4 &= A\ (3,6)\,(1,2)\,(1,4)\,(1,5)\,(2,4)\,(2,5)\,(4,5)\,(3,7)\\
{\theta_{26}}^4 &= A\ (2,6)\,(4,6)\,(1,3)\,(1,5)\,(3,5)\,(2,4)\,(2,7)\,(4,7)\\
{\theta_{27}}^4 &= A\ (1,6)\,(5,6)\,(1,5)\,(2,3)\,(2,4)\,(3,4)\,(2,7)\,(3,7)\,(4,7)\\
{\theta_{28}}^4 &= A\ (1,6)\,(3,6)\,(1,3)\,(2,4)\,(2,5)\,(4,5)\,(2,7)\,(4,7)\,(5,7)\\
{\theta_{29}}^4 &= A\ (1,6)\,(2,6)\,(4,6)\,(3,5)\,(1,2)\,(1,4)\,(2,4)\,(3,7)\,(5,7)\\
{\theta_{30}}^4 &= A\ (5,6)\,(1,2)\,(1,3)\,(1,4)\,(2,3)\,(2,4)\,(3,4)\,(5,7)\\
{\theta_{31}}^4 &= A\ (1,6)\,(2,6)\,(3,6)\,(1,2)\,(1,3)\,(2,3)\,(4,5)\,(4,7)\,(5,7)
\end{aligned}$$

$$\theta_{32}{}^4 = A\ (1,6)\,(4,6)\,(2,3)\,(2,5)\,(3,5)\,(1,4)\,(2,7)\,(3,7)\,(5,7)$$
$$\theta_{33}{}^4 = A\ (2,6)\,(5,6)\,(1,3)\,(1,4)\,(3,4)\,(2,5)\,(2,7)\,(5,7)$$
$$\theta_{34}{}^4 = A\ (2,6)\,(3,6)\,(1,4)\,(1,5)\,(4,5)\,(2,3)\,(2,7)\,(3,7)$$
$$\theta_{35}{}^4 = A\ (4,6)\,(1,2)\,(1,3)\,(1,5)\,(2,3)\,(2,5)\,(3,5)\,(4,7)$$
$$\theta_{36}{}^4 = A\ (1,6)\,(2,6)\,(5,6)\,(1,2)\,(1,5)\,(2,5)\,(3,4)\,(3,7)\,(4,7)$$

By using the set of equations given above we have several choices for $a_1, \cdots, a_5$ in terms of theta constants.

Branch Points	Possible Ratios		
a_1^2	$\left(\frac{\theta_{36}^2\theta_{22}^2}{\theta_{33}^2\theta_{19}^2}\right)^2$	$\left(\frac{\theta_{31}^2\theta_{21}^2}{\theta_{34}^2\theta_{24}^2}\right)^2$	$\left(\frac{\theta_{29}^2\theta_{1}^2}{\theta_{26}^2\theta_{2}^2}\right)^2$
a_2^2	$\left(\frac{\theta_{4}^2\theta_{29}^2}{\theta_{2}^2\theta_{17}^2}\right)^2$	$\left(\frac{\theta_{36}^2\theta_{7}^2}{\theta_{15}^2\theta_{19}^2}\right)^2$	$\left(\frac{\theta_{31}^2\theta_{13}^2}{\theta_{9}^2\theta_{24}^2}\right)^2$
a_3^2	$\left(\frac{\theta_{4}^2\theta_{22}^2}{\theta_{33}^2\theta_{17}^2}\right)^2$	$\left(\frac{\theta_{11}^2\theta_{31}^2}{\theta_{24}^2\theta_{6}^2}\right)^2$	$\left(\frac{\theta_{7}^2\theta_{1}^2}{\theta_{26}^2\theta_{15}^2}\right)^2$
a_4^2	$\left(\frac{\theta_{11}^2\theta_{29}^2}{\theta_{2}^2\theta_{6}^2}\right)^2$	$\left(\frac{\theta_{21}^2\theta_{7}^2}{\theta_{15}^2\theta_{34}^2}\right)^2$	$\left(\frac{\theta_{22}^2\theta_{13}^2}{\theta_{9}^2\theta_{33}^2}\right)^2$
a_5^2	$\left(\frac{\theta_{4}^2\theta_{21}^2}{\theta_{34}^2\theta_{17}^2}\right)^2$	$\left(\frac{\theta_{11}^2\theta_{36}^2}{\theta_{19}^2\theta_{6}^2}\right)^2$	$\left(\frac{\theta_{13}^2\theta_{1}^2}{\theta_{26}^2\theta_{9}^2}\right)^2$

Let's select the following choices for $a_1, \cdots, a_5$.

$$a_1 = \frac{\theta_{31}^2\theta_{21}^2}{\theta_{34}^2\theta_{24}^2},\ a_2 = \frac{\theta_{31}^2\theta_{13}^2}{\theta_{9}^2\theta_{24}^2},\ a_3 = \frac{\theta_{11}^2\theta_{31}^2}{\theta_{24}^2\theta_{6}^2},\ a_4 = \frac{\theta_{21}^2\theta_{7}^2}{\theta_{15}^2\theta_{34}^2},\quad a_5 = \frac{\theta_{13}^2\theta_{1}^2}{\theta_{26}^2\theta_{9}^2}.$$

This completes the proof. □

Remark 18. *Unlike the genus 2 case, here only* θ_1, θ_6, θ_7, θ_{11}, θ_{15}, θ_{24}, θ_{31} *are from one of the Göpel groups.*

4.1.1. *Genus 3 non-hyperelliptic cyclic curves.* Using the Thomae's like formula for cyclic curves, for each cyclic curve $y^n = f(x)$ one can express the roots of $f(x)$ in terms of ratios of theta functions as in the hyperelliptic case. In this section we study such curves for $g = 3$. We only consider the families of curves with positive dimension since the curves which belong to 0-dimensional families are well known. The proof of the following lemma can be found in [12].

Lemma 19. *Let* f *be a meromorphic function on* $\mathcal{X}$*, and let*

$$(f) = \sum_{i=1}^{m} b_i - \sum_{i=1}^{m} c_i$$

be the divisor defined by f*. Let's take paths from* P_0 *(initial point) to* b_i *and* P_0 *to* c_i *so that* $\sum_{i=1}^m \int_{P_0}^{b_i} \omega = \sum_{i=1}^m \int_{P_0}^{c_i} \omega$.

For an effective divisor $P_1 + \cdots + P_g$ *we have*

$$f(P_1)\cdots f(P_g) = \frac{1}{E}\prod_{k=1} \frac{\theta(\sum_i \int_{P_0}^{P_i} \omega - \int_{P_0}^{b_k} \omega - \triangle, \tau)}{\theta(\sum_i \int_{P_0}^{P_i} \omega - \int_{P_0}^{c_k} \omega - \triangle, \tau)} \tag{15}$$

where E *is a constant independent of* $P_1, \ldots, P_g$*, the integrals from* P_0 *to* P_i *take the same paths both in the numerator and in the denominator,* $\triangle$ *denotes the Riemann's constant and* $\int_{P_0}^{P_i} \omega = \left(\int_{P_0}^{P_i} \omega_1, \ldots, \int_{P_0}^{P_i} \omega_g\right)^t$.

Notice that the definition of thetanulls is different in this part from the definitions of the hyperelliptic case. We define the following three theta constants.

$$\theta_1 = \theta \begin{bmatrix} 0 & \frac{1}{6} & 0 \\ \frac{2}{3} & \frac{1}{6} & \frac{2}{3} \end{bmatrix} \quad \theta_2 = \theta \begin{bmatrix} 0 & \frac{1}{6} & 0 \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{3} \end{bmatrix} \quad \theta_3 = \theta \begin{bmatrix} 0 & \frac{1}{6} & 0 \\ 0 & \frac{1}{6} & 0 \end{bmatrix}$$

Next we consider the cases 16, 18, 19 from Table 4.

Case 18: If the automorphism group is C_3 then the equation of $\mathcal{X}$ is given by

$$y^3 = x(x-1)(x-s)(x-t).$$

Let Q_i where $i = 1..5$ be ramifying points in the fiber of $0, 1, s, t, \infty$ respectively. Consider the meromorphic function $f = x$ on $\mathcal{X}$ of order 3. Then we have $(f) = 3Q_1 - 3Q_5$. By applying the Lemma 19 with $P_0 = Q_5$ and an effective divisor $2Q_2 + Q_3$ we have the following.

$$Es = \prod_{k=1}^{3} \frac{\theta(2\int_{Q_5}^{Q_2}\omega + \int_{Q_5}^{Q_3}\omega - \int_{Q_5}^{b_k}\omega - \triangle, \tau)}{\theta(2\int_{Q_5}^{Q_2}\omega + \int_{Q_5}^{Q_3}\omega - \triangle, \tau)} \tag{16}$$

Again apply the Lemma 19 with an effective divisor Q_2+2Q_3 we have the following.

$$Es^2 = \prod_{k=1}^{3} \frac{\theta(\int_{Q_5}^{Q_2}\omega + 2\int_{Q_5}^{Q_3}\omega - \int_{Q_5}^{b_k}\omega - \triangle, \tau)}{\theta(\int_{Q_5}^{Q_2}\omega + 2\int_{Q_5}^{Q_3}\omega - \triangle, \tau)} \tag{17}$$

By dividing Eq. (17) by Eq. (16) we have,

$$\begin{aligned} s = &\prod_{k=1}^{3} \frac{\theta(\int_{Q_5}^{Q_2}\omega + 2\int_{Q_5}^{Q_3}\omega - \int_{Q_5}^{b_k}\omega - \triangle, \tau)}{\theta(\int_{Q_5}^{Q_2}\omega + 2\int_{Q_5}^{Q_3}\omega - \triangle, \tau)} \\ &\times \prod_{k=1}^{3} \frac{\theta(2\int_{Q_5}^{Q_2}\omega + \int_{Q_5}^{Q_3}\omega - \triangle, \tau)}{\theta(2\int_{Q_5}^{Q_2}\omega + \int_{Q_5}^{Q_3}\omega - \int_{Q_5}^{b_k}\omega - \triangle, \tau)} \end{aligned} \tag{18}$$

By a similar argument we have

$$\begin{aligned} t = &\prod_{k=1}^{3} \frac{\theta(\int_{Q_5}^{Q_2}\omega + 2\int_{Q_5}^{Q_4}\omega - \int_{Q_5}^{b_k}\omega - \triangle, \tau)}{\theta(\int_{Q_5}^{Q_2}\omega + 2\int_{Q_5}^{Q_4}\omega - \triangle, \tau)} \\ &\times \prod_{k=1}^{3} \frac{\theta(2\int_{Q_5}^{Q_2}\omega + \int_{Q_5}^{Q_4}\omega - \triangle, \tau)}{\theta(2\int_{Q_5}^{Q_2}\omega + \int_{Q_5}^{Q_4}\omega - \int_{Q_5}^{b_k}\omega - \triangle, \tau)} \end{aligned} \tag{19}$$

Computing the right hand side of Eq. (18) and Eq. (19) was the one of the main points of [11]. Finally, we have

$$s = \frac{\theta_2^3}{\theta_1^3}, \ \textit{and} \ \ r = \frac{\theta_3^3}{\theta_1^3}.$$

Case 19: If the group is C_6 then the equation is $y^3 = x(x-1)(x-s)(x-t)$ with $s = 1 - t$. By using results from case 18, we have

$$\theta_2^3 = \theta_1^3 - \theta_3^3.$$

Case 16: In this case the equation of $\mathcal{X}$ is given by

$$y^4 = x(x-1)(x-t).$$

This curve has 4 ramifying points Q_i where $i = 1..4$ in the fiber of $0, 1, t, \infty$ respectively. Consider the meromorphic function $f = x$ on $\mathcal{X}$ of order 4. Then we have $(f) = 4Q_1 - 4Q_4$. By applying the Lemma 19 with $P_0 = Q_4$ and an effective divisor $2Q_2 + Q_3$ we have the following.

$$Et = \prod_{k=1}^{4} \frac{\theta(2\int_{Q_4}^{Q_2}\omega + \int_{Q_4}^{Q_3}\omega - \int_{Q_4}^{b_k}\omega - \triangle, \tau)}{\theta(2\int_{Q_4}^{Q_2}\omega + \int_{Q_4}^{Q_3}\omega - \triangle, \tau)} \tag{20}$$

Again apply the Lemma 19 with an effective divisor Q_2+2Q_3 we have the following.

$$Et^2 = \prod_{k=1}^{4} \frac{\theta(\int_{Q_4}^{Q_2}\omega + 2\int_{Q_4}^{Q_3}\omega - \int_{Q_4}^{b_k}\omega - \triangle, \tau)}{\theta(\int_{Q_4}^{Q_2}\omega + 2\int_{Q_4}^{Q_3}\omega - \triangle, \tau)} \tag{21}$$

We have the following by dividing Eq. (21) by Eq. (20)

$$\begin{aligned} t = &\prod_{k=1}^{4} \frac{\theta(\int_{Q_4}^{Q_2}\omega + 2\int_{Q_4}^{Q_3}\omega - \int_{Q_4}^{b_k}\omega - \triangle, \tau)}{\theta(\int_{Q_4}^{Q_2}\omega + 2\int_{Q_4}^{Q_3}\omega - \triangle, \tau)} \\ &\times \prod_{k=1}^{4} \frac{\theta(2\int_{Q_4}^{Q_2}\omega + \int_{Q_4}^{Q_3}\omega - \triangle, \tau)}{\theta(2\int_{Q_4}^{Q_2}\omega + \int_{Q_4}^{Q_3}\omega - \int_{Q_4}^{b_k}\omega - \triangle, \tau)} \end{aligned} \tag{22}$$

In order to compute the explicit formula for t one has to find the integrals on the right hand side. Such computations are long and tedious and we intend to include them in further work.

Remark 20. *In the case 16) of Table 4, the parameter t is given by*

$$\theta[e]^4 = A(t-1)^4 t^2,$$

where $[e]$ is the theta characteristics corresponding to the partition $(\{1\},\{2\},\{3\},\{4\})$ and A is a constant; see [8] *for details. However, this is not satisfactory since we would like t as a rational function in terms of theta. The methods in* [8] *do not lead to another relation among t and the thetanulls since the only partition we could take is the above.*

Summarizing all of the above we have:

Lemma 21. *Let $\mathcal{X}$ be a non-hyperelliptic genus 3 curve. The following are true:*

i): *If $Aut(\mathcal{X}) \cong C_3$, then $\mathcal{X}$ is isomorphic to a curve with equation*

$$y^3 = x(x-1)\left(x - \frac{\theta_2^3}{\theta_1^3}\right)\left(x - \frac{\theta_3^3}{\theta_1^3}\right).$$

ii): *If $Aut(\mathcal{X}) \cong C_6$, then $\mathcal{X}$ is isomorphic to a curve with equation*

$$y^3 = x(x-1)\left(x - \frac{\theta_2^3}{\theta_1^3}\right)\left(x - \frac{\theta_3^3}{\theta_1^3}\right) \text{ with } \theta_2^3 = \theta_1^3 - \theta_3^3.$$

iii): *If $Aut(\mathcal{X})$ is isomorphic to the group with GAP identity $(16, 13)$, then $\mathcal{X}$ is isomorphic to a curve with equation*

$$y^4 = x(x-1)(x-t) \text{ with}$$

where t is given by Eq. (22).

It seems possible to generalize such techniques of computing the branch points in terms of the theta functions for any cyclic cover of the projective line. We intend to pursue the ideas of these papers in further work.

Acknowledgements: The first ideas of this paper started during a visit of the second and third author at Boston University during the Summer 2006. Both the second and third author want to thank Prof. Previato for making that visit possible.

References

[1] G. Cardona, J. Quer, Field of moduli and field of definition for curves of genus 2. Computational aspects of algebraic curves, 71–83, Lecture Notes Ser. Comput., 13, World Sci. Publ., Hackensack, NJ, 2005.
[2] A. Krazer, Lehrbuch der Thetafunctionen, Chelsea, New York, (1970).
[3] R. Kuhn, Curves of genus 2 with split Jacobian. Trans. Amer. Math. Soc. 307 (1988), no. 1, 41–49.
[4] H.F. Baker, Abelian Function, Abel's theorem and the allied theory of theta functions, (1897).
[5] K. Magaard, T. Shaska, S. Shpectorov, H. Vlklein, The locus of curves with prescribed automorphism group. Communications in arithmetic fundamental groups (Kyoto, 1999/2001). Sūrikaisekikenkyūsho Kōkyūroku No. 1267 (2002), 112–141.
[6] D. Mumford, Tata lectures on theta. II. Jacobian theta functions and differential equations. With the collaboration of C. Musili, M. Nori, E. Previato, M. Stillman and H. Umemura. Progress in Mathematics, 43. Birkhuser Boston, Inc., Boston, MA, 1984.
[7] D. Mumford, Tata lectures on theta. I. With the assistance of C. Musili, M. Nori, E. Previato and M. Stillman. Progress in Mathematics, 28. Birkhuser Boston, Inc., Boston, MA, 1983. xiii+235 pp.
[8] A.Nakayashiki, On the Thomae formula for Z_N curves, *Publ. Res. Inst. Math. Sci.*, vol 33 (1997), no. 6, pg. 987–1015.
[9] T. Shaska, Curves of genus 2 with (N, N) decomposable Jacobians, *J. Symbolic Comput.*, vol. 31, Nr. 5, 2001, 603–617.
[10] *Algebraic curves and their applications* `http://www.albmath.org/algcurves/`
[11] H. Shiga, On the representation of the Picard modular function by θ constants. I, II., *Publ. Res. Inst. Math. Sci.*, vol. 24, (1988), no. 3, pg. 311–360.
[12] H.E. Rauch and H.M.Farkas, Theta functions with applications to Riemann surfaces, Williams and Wilkins, Baltimore, 1974.

Department of Mathematics and Statistics,
Boston University,
Boston, MA 02215-2411
Email: ep@bu.edu

Department of Mathematics and Statistics,
Oakland University,
546 Science and Eng. Building,
Rochester, MI 48309
Email: shaska@oakland.edu

Department of Mathematics and Statistics,
Oakland University,
389 Science and Eng. Build.,
Rochester, MI 48309
Email: gswijesi@oakland.edu

ALBANIAN JOURNAL
OF MATHEMATICS
Volume 1, Number 4, Pages 271–282
ISSN 1930-1235: (2007)

MODULAR EQUATIONS OF ORDER p AND THETA FUNCTIONS

YAACOV KOPELIOVICH

Dedicated to Mike Fried on his 65-th birthday
for constant mathematical inspiration.

ABSTRACT. Let p be a prime integer and $\mathbf{H_g}$ be a collection of complex positive definite symmetric $g \times g$ matrices τ. Denote by $p\tau$ the multiplication of τ by p. In this note we describe an explicit process to obtain algebraic identities between theta functions with integral characteristics evaluated at τ and $p\tau$. For $g = 1$ this produces modular equations between $\lambda(\tau), \lambda(p\tau)$ where $\lambda(\tau)$ is the invariant associated with elliptic curve generated by τ, described by the equation: $y^2 = x(x-1)(x-\lambda(\tau_1))$. Consequently, if $g > 1$ the algebraic identities we obtain might serve as a higher dimensional generalization for the one dimensional modular equations.

1. INTRODUCTION

Let τ_1 be a complex number such that $Im(\tau_1) > 0$ and Z_{τ_1} is the lattice generated by $\{1, \tau_1\}$ in $\mathbb{C}$. Let $C_1 = \mathbb{C}/Z_{\tau_1}$, be the corresponding analytic elliptic curve. The algebraic equation of this curve is

$$y^2 = x(x-1)(x-\lambda(\tau_1)). \tag{1}$$

and $\lambda(\tau_1)$ is the invariant corresponding to τ_1 in this equation.

Definition 1.1. Let p be a prime number. A modular equation of order p for λ, is an algebraic equation between $\lambda(\tau_1)$ and $\lambda(p\tau_1)$.

These equations appeared naturally in the theory of elliptic integrals, because they describe algebraically the analytical multiplication by p on the lattice Z_{τ_1}. In equivalent, but more contemporary terms, this equation describes the λ-invariant of curves C_2, where $\phi : C_1 \mapsto C_2$ is an isogeny (finite homomorphism) of order p. Equations of this type have an important role in Galois theory of complex multiplications, since if C_1 is a curve with complex multiplication then $\lambda(p\tau_1)$ generates interesting field extension of $Q^{ab}(\tau_1)$. Another application of one dimensional modular equations is algorithms for rapid calculation of π, [Bo].

In recent years, there are applications of modular equations to point counting algorithms of elliptic curve above finite fields F_{p^i}. Modular equations of order p are used to compute explicit canonical lifting of elliptic curves above finite fields F_{p^i}, of characteristics p to the corresponding p - adic field above it. Using the lifting we calculate the trace of the Frobenius operator on the p- adic field. Applying fixed

2000 *Mathematics Subject Classification.* 14K25,32G20.
Key words and phrases. Theta functions, modular equations, λ function .

point formulas, we find the number of points of elliptic curves above finite fields F_{p^i}, [Ma] explains the general framework for this type of algorithms.

For $p = 2$ the modular equation of level 2 is the arithmetic geometric mean (AGM) in disguise. If $a_0 = a, b_0 = b$ are real numbers we define the AGM iteration as:

$$a_n = \frac{a_n + b_n}{2} \tag{2}$$

$$b_n = \sqrt{a_n b_n} \tag{3}$$

This iteration has a strong link with the modular equations of order 2. These sequences converge to a common limit denoted by $AGM(a, b)$ see [Bo]. Mestre [Me] used the AGM iteration to suggest an algorithm for point counting defined over F_{2^i}. In [Me], Mestre uses a higher dimensional analogue of the AGM and produces an algorithm to count the number of points of hyperelliptic curves above fields of characteristics 2. In a recent work, Lubicz, Carls, and Kohel [CKL] generalized Mestre's method further in a different direction. They construct curves that have good cryptographic properties using an iteration of order 3. In one dimensional case their iteration seems to be closely related to modular equations of order 3. These results motivate the question whether there exists a theory of modular equations for τ an element of $\mathbf{H_g}$. To suggest a possible generalization recall that $\lambda(\tau_1)$ is a quotient of theta functions that is:

$$\lambda(\tau_1) = \frac{\Theta^4 \begin{bmatrix} 0 \\ 1 \end{bmatrix} (0, \tau_1)}{\Theta^4 \begin{bmatrix} 0 \\ 0 \end{bmatrix} (0, \tau_1)}.$$

These are analytic functions that we define in the first section of the paper. $\lambda(p\tau_1)$ is the same quotient evaluated at $p\tau_1$. Hence, modular equations become identities between theta functions evaluated at τ_1 and those evaluated at $p\tau_1$. While an elementary dimension argument shows that analogues of $\lambda(\tau_1)$ do not exist for general $\tau \in \mathbf{H_g}$ theta functions do. Thus, the question of modular equations is reduced to finding a way to produce certain identities between higher dimensional theta functions evaluated at τ and $p\tau$. In this note, we apply methods from [Ko] to suggest such a procedure. The procedure constructs equations between

$$\Theta \begin{bmatrix} \eta_i \\ \epsilon_i \end{bmatrix} (0, \tau) \ \ and \ \Theta \begin{bmatrix} \eta_i \\ \epsilon_i \end{bmatrix} (0, p\tau),$$

for any $p > 2$ and η_i, ϵ_i are g integral characteristics. We stress that in addition to applications similar to the one dimensional case and possibly in cryptography we believe that an existence of such a procedure should lead to other applications that are not present in the 1-dimensional case. For example such a procedure should produce algebraic conditions that characterize Abelian varieties that are isogenous to multiplication of elliptic curves. The problem treated in this note is addressed in the literature with different approaches. Modular equations for hyperelliptic curves through l-torsion subgroups of Jacobians were defined and treated in [GS]. [CKL] treated the case for $p = 3$ and [CL] seem to treat the general p along the same lines. Their work produces identities that use the theory of algebraic theta functions and

Riemann's theta formula. The relation of these identities to the current work is not obvious and we plan to investigate it further in the future.

We review the structure of this note: The first section explains the process to obtain identities between theta functions with integral characteristics at τ and $p\tau$. We apply these results in the second section to the one dimensional case and explain how this leads to a proof of existence of modular equations to the λ function. This serves as an alternative to the classical theory of modular polynomials that has no analogue in the higher dimensional case. In the last section we produce modular equations for $p = 3$ and 7.

Acknowledgements: We thank David Lubicz whose interest in these questions prompted the initial motivation for this work. We also thank David Kohel and Hershel Farkas for reading and providing valuable suggestions on an earlier version for this note. We especially thank the referee for very constructive remarks that substantially improved the presentation of this note. This work was partially done while the author visited Oakland University and the author thanks the Department of Mathematics and the algebra group for their support and hospitality.

2. Modular equations

We remind the reader the definition and main properties of theta functions:

Definition 2.1. Let τ be a complex $g \times g$ matrix such that:

- $\tau = \tau^t$ i.e. τ is symmetric
- $Im\tau$ is a positive definite quadratic form.

Then, $\tau \in \mathbf{H_g}$. Let $\begin{bmatrix} \varepsilon \\ \varepsilon' \end{bmatrix}$ be a real $2g$ vector. **Theta function** is a complex analytic function on $C^g \times \mathbf{H_g}$ such that:

$$\Theta \begin{bmatrix} \varepsilon \\ \varepsilon' \end{bmatrix} (z,\tau) = \sum_{l \in \mathbb{Z}^g} exp 2\pi i \left\{ \frac{1}{2} \left(l + \frac{\varepsilon}{2} \right)^t \tau \left(l + \frac{\varepsilon}{2} \right) + \left(l + \frac{\epsilon}{2} \right)^t \left(z + \frac{\epsilon'}{2} \right) \right\}$$

Note that the definition is the classical definition of theta function. The modern authors omit the factor $\frac{1}{2}$. We list the main properties of theta functions:

- $\Theta \begin{bmatrix} \varepsilon + 2m \\ \varepsilon' + 2e \end{bmatrix} (z,\tau) = exp\pi i \left\{ \varepsilon^t e \right\} \Theta \begin{bmatrix} \varepsilon \\ \varepsilon' \end{bmatrix} (z,\tau)$ and $m, e \in Z^g$
- $\Theta \begin{bmatrix} \varepsilon \\ -\varepsilon' \end{bmatrix} (z,\tau) = \Theta \begin{bmatrix} \varepsilon \\ \varepsilon' \end{bmatrix} (-z,\tau)$
- $\Theta \begin{bmatrix} \varepsilon \\ \varepsilon' \end{bmatrix} (z+n+m^t\tau,\tau) = \exp 2\pi i \left\{ \frac{n^t\varepsilon - m^t\varepsilon'}{2} - m^t z - m^t \tau m \right\} \Theta \begin{bmatrix} \varepsilon \\ \varepsilon' \end{bmatrix} (z,\tau)$

For the proof of these properties that follow from a careful series manipulation see [Mu], [RF] or [Ko]. We remind the reader the notion of integral (rational) theta characteristics:

Definition 2.2. The functions

$$\Theta \begin{bmatrix} \varepsilon \\ \varepsilon' \end{bmatrix} (z,\tau)$$

are called theta functions with **integral** (**rational**) characteristics if $\begin{bmatrix} \varepsilon \\ \varepsilon' \end{bmatrix} \in \mathbb{Z}^{2g} \left(\mathbb{Q}^{2g}\right)$

From property (2) we see that we can assume that $\begin{bmatrix} \varepsilon \\ \varepsilon' \end{bmatrix} \in Z_2^{2g}$ Further $\Theta \begin{bmatrix} \varepsilon \\ \varepsilon' \end{bmatrix} (z,\tau)$ will be even or odd if the scalar product $\varepsilon'^t\varepsilon = 0,1$ respectively. This motivates the following definition:

Definition 2.3. The integral characteristics $\begin{bmatrix} \varepsilon \\ \varepsilon' \end{bmatrix}$ is called **even** (**odd**) if $\varepsilon'^t\varepsilon = 0,1$

Let us cite the duplication formula for higher dimensional theta functions that will be important in the sequel:

$$\Theta^2 \begin{bmatrix} \epsilon \\ \epsilon_1 \end{bmatrix} (0,\tau) = \sum_{\alpha' \in \mathbb{Z}_2^g} \Theta \begin{bmatrix} \epsilon+\alpha' \\ 2\epsilon_1 \end{bmatrix} (0,2\tau) \Theta \begin{bmatrix} \alpha' \\ 0 \end{bmatrix} (0,2\tau)$$

Here $\epsilon, \epsilon_1 \in \mathbb{Q}^g$ are any g rational characteristics. The proof is in [Mu] or [RF]. In particular, replace τ with 2τ on both sides of the equation to obtain that:

$$\Theta^2 \begin{bmatrix} \epsilon \\ \epsilon_1 \end{bmatrix} (0,2\tau) = \sum_{\alpha' \in \mathbb{Z}_2^g} \Theta \begin{bmatrix} \epsilon+\alpha' \\ 2\epsilon_1 \end{bmatrix} (0,4\tau) \Theta \begin{bmatrix} \alpha' \\ 0 \end{bmatrix} (0,4\tau) \tag{4}$$

Let us show the following lemma:

Lemma 2.4. *Let $\delta \in \mathbb{Z}_2^g$ then:*

$$\Theta \begin{bmatrix} 0 \\ \delta \end{bmatrix} (0,\tau) = \sum_{\beta \in \mathbb{Z}_2^g} exp\left(\pi i \delta \cdot \beta^t\right) \Theta \begin{bmatrix} \beta \\ 0 \end{bmatrix} (0,4\tau)$$

Proof. We write by the definition of theta functions:

$$\Theta \begin{bmatrix} 0 \\ \delta \end{bmatrix} (0,\tau) = \sum_{l \in \mathbb{Z}^g} exp 2\pi i \left\{ \frac{1}{2} l\tau l^t + l^t \frac{\delta}{2} \right\}.$$

Rewrite the right hand side of the last equation as:

$$\sum_{m \in Z^g, \beta \in Z_2^g} exp 2\pi i \left\{ \frac{1}{2} (2m+\beta)\tau (2m+\beta)^t + (2m+\beta)^t \frac{\delta}{2} \right\}$$

Since $exp(2\pi i m\delta) = 1$ this equals to

$$\sum_{m \in \mathbb{Z}^g, \beta \in Z_2^g} \exp(\pi i \delta^t \beta) \times \exp 2\pi i \left\{ \frac{1}{2} \left(m + \frac{\beta}{2}\right) 4\tau \left(m + \frac{\beta}{2}\right)^t \right\}$$

then the last sum equals to

$$\sum_{\beta \in \mathbb{Z}_2^g} exp\left(\pi i \delta \cdot \beta^t\right) \Theta \begin{bmatrix} \beta \\ 0 \end{bmatrix} (0,4\tau),$$

by the definition of theta functions.

□

Corollary 2.5. *The following identity holds*

$$\Theta\begin{bmatrix}\beta\\0\end{bmatrix}(0,4\tau)=\frac{1}{2^g}\sum_{\delta\in Z_2^g}\exp\left(-\pi i\delta^t\cdot\beta\right)\Theta\begin{bmatrix}0\\\delta\end{bmatrix}(0,\tau)$$

Proof. We can treat $\Theta\begin{bmatrix}\beta\\0\end{bmatrix}(0,4\tau)$ as unknowns in a system of linear equations in which the elements of the $2^g\times 2^g$ matrix A are $exp(\pi\delta^t\beta)$. Multiply A by its Hermitian transpose A^*. Then $(AA^*)_{ii}=2^g$ since the diagonal element equals

$$\sum_{\delta}\exp(\pi i\delta^t\beta)\exp(-\pi i\delta^t\beta)=2^g.$$

If $i\neq j$ then $(AA^*)_{ij}=0$. This is because this element can be regarded as a sum of a nontrivial character on the group Z_2^g.

□

We apply the formula to Eq. (4). If ϵ_1 is an integer we rewrite the equation as:

(5)
$$\Theta^2\begin{bmatrix}\epsilon\\\epsilon_1\end{bmatrix}(0,2\tau)=\sum_{\alpha'\in\mathbb{Z}_2^g}exp\left(\pi i(\epsilon+\alpha')\epsilon_1\right)\Theta\begin{bmatrix}\epsilon+\alpha'\\0\end{bmatrix}(0,4\tau)\Theta\begin{bmatrix}\alpha'\\0\end{bmatrix}(0,4\tau)$$

Applying Corollary **2.4** the last equation equals to

$$\sum_{\alpha,\gamma,\delta\in\mathbb{Z}_2^g}\frac{1}{2^{2g}}c_{\epsilon,\epsilon',\alpha,\gamma,\delta}\Theta\begin{bmatrix}0\\\gamma\end{bmatrix}(0,\tau)\Theta\begin{bmatrix}0\\\delta\end{bmatrix}(0,\tau)\tag{6}$$

and

$$c_{\epsilon,\epsilon',\alpha',\gamma,\delta}=\exp\left(\pi i(\epsilon+\alpha')^t\epsilon_1\right)\times\exp\left(\pi i(\epsilon'+\alpha')^t\gamma\right)\times\exp\left(\pi i\alpha'^t\delta\right)$$

Note that for fixed δ,γ the total coefficient of the product $\Theta\begin{bmatrix}0\\\gamma\end{bmatrix}(0,\tau)\Theta\begin{bmatrix}0\\\delta\end{bmatrix}(0,\tau)$ is:

$$\sum_{\alpha}\exp\left(\pi i(\epsilon+\alpha')^t\epsilon_1\right)\times\exp\left(\pi i(\epsilon'+\alpha')^t\gamma\right)\times\exp\left(\pi i\alpha'^t\delta\right)$$

which equals 0 unless $\epsilon_1+\gamma+\delta=0$. In the latter case the coefficient is $2^g\exp(\pi i\epsilon^t\delta)$. Summarizing, we obtain the following corollary:

Corollary 2.6.

$$\Theta^2\begin{bmatrix}\epsilon\\\epsilon_1\end{bmatrix}^2(0,2\tau)=\sum_{\delta\in Z_2^g}\exp\left(\pi i\epsilon^t\delta\right)\Theta\begin{bmatrix}0\\\delta\end{bmatrix}(0,\tau)\Theta\begin{bmatrix}0\\\epsilon_1-\delta\end{bmatrix}(0,\tau)\tag{7}$$

We apply the last corollary to explain a procedure to obtain a higher dimensional modular equations analogues for any prime p.

Definition 2.7. Let D_n be a set of vectors $\begin{bmatrix}\epsilon\\\epsilon_1\end{bmatrix}$, such that

- $\epsilon\in\mathbb{Z}^g,\epsilon_i=0,1.$
- $\epsilon_1\in\mathbb{Q}^g,\epsilon_{1i}=\frac{l}{2^n},0\leq l<2^n$

For each $\tau \in \mathbf{H_g}$ let $\alpha_n(\tau) = \Theta \left[\begin{array}{c} \epsilon \\ \epsilon_1 \end{array} \right](0,\tau)$ and $\left[\begin{array}{c} \epsilon \\ \epsilon_1 \end{array} \right] \in D_n$. $\tau \mapsto \alpha_n(\tau)$ induces a map from $\psi_n(\tau) : \mathbf{H_g} \mapsto CP^l$ (l - number of vectors in D_n.)

Let $X_n = \psi_n(\mathbf{H_g})$. Then X_1 is the image in $CP^{2^{2g}-1}$ where ϵ, ϵ_1 are integral characteristics. There is a map $\phi_n : X_n \mapsto X_1$ which omits the non integer characteristics in the definition of $\psi_n(\tau)$.

Lemma 2.8. *The map ϕ_n is a finite map from $X_n \mapsto X_1$.*

Proof. Because of the transformation formula for theta functions [RF] under the action of $Sp(g,\mathbf{Z})$, there exists a subgroup Δ_n of finite index in $Sp(g,\mathbf{Z})$ such that $\psi_n(\tau)$ induces a map $\beta_n(\tau)$, $\beta_n : \mathbf{H_g}/\Delta_n \mapsto X_n$ (Note: Δ_n **is not a congruence subgroup of level** n.) Hence, the map ϕ_n factors through a map

$$\bar{\phi_n} : \mathbf{H_g}/\Delta_n \mapsto \mathbf{H_g}/\Delta_1,$$

which is clearly finite since Δ_n has a finite index inside $Sp(g,\mathbf{Z})$.

□

Before stating the theorem that describes the map ϕ_n more explicitly we state the following definition:

Definition 2.9. Let H be a complex analytic domain and $f_1...f_n : H \mapsto C$ be complex analytic functions. We call f constructible from $f_1...f_n$ if

- f is algebraic above $\mathbb{C}(f_1...f_n)$
- The Galois group of $C(f)$ above $f_1...f_n$ is solvable. equivalently f can be expressed through radical expressions involving $f_1...f_n$.

Theorem 2.10. *Let $\epsilon_1 \in \mathbb{Q}^g, \epsilon_{1i} = \frac{l}{2^n}$ and $\epsilon \in Z_2^g$. Then, $\Theta \left[\begin{array}{c} \epsilon \\ \epsilon_1 \end{array} \right](0,\tau)$ is constructible from $\Theta \left[\begin{array}{c} \eta \\ \eta_1 \end{array} \right](0,\tau)$ and η, η_1 are integral characteristics.*

Proof. Assume inductively that the theorem is true for all characteristics $\Theta \left[\begin{array}{c} \delta \\ \delta_1 \end{array} \right](0,\tau)$ such that $\delta \in \mathbb{Z}^g$ and $\delta_{1i} = \frac{l}{2^{n-1}}$. The duplication formula implies:

$$\Theta^2 \left[\begin{array}{c} \epsilon \\ \epsilon_1 \end{array} \right](0,\tau) = \sum_{\alpha' \in \mathbb{Z}_2^g} \Theta \left[\begin{array}{c} \epsilon + \alpha' \\ 2\epsilon_1 \end{array} \right](0,2\tau) \Theta \left[\begin{array}{c} \alpha' \\ 0 \end{array} \right](0,2\tau)$$

But

$$\Theta \left[\begin{array}{c} \epsilon + \alpha' \\ 2\epsilon_1 \end{array} \right](0,2\tau)$$

satisfies the induction hypothesis. So it is constructible from

$$\Theta^2 \left[\begin{array}{c} \eta \\ \eta_1 \end{array} \right](0,2\tau)$$

Apply the formulas from Corollary **2.6** to see that

$$\Theta^2 \left[\begin{array}{c} \eta \\ \eta_1 \end{array} \right](0,2\tau)$$

is constructible from $\Theta \left[\begin{array}{c} \eta \\ \eta_1 \end{array} \right](0,\tau)$. Hence, $\Theta \left[\begin{array}{c} \epsilon \\ \epsilon_1 \end{array} \right](0,\tau)$ is constructible. □

Note that the proof gives slightly more i.e. a recursive process how to construct the expressions

$$\Theta \left[\begin{array}{c} \epsilon \\ \epsilon_1 \end{array} \right] (0,\tau)$$

from

$$\Theta \left[\begin{array}{c} \eta \\ \eta_1 \end{array} \right] (0,\tau)$$

We reformulate it in the following corollary:

Corollary 2.11. *Let $P \in X_1$ and let $Q \in \phi_n^{-1}(P)$ then there exists a process that produces an algebraic relationship between the coordinates of Q and coordinates of P.*

We rely on the last theorem and the methods developed in [Ko] to obtain process to produce generalized modular equation in any dimension.

Definition 2.12. Let p be a prime. A modular equation of order p will be any non trivial polynomial identity between $\Theta \left[\begin{array}{c} \eta \\ \eta_1 \end{array} \right] (0,\tau)$ and $\Theta \left[\begin{array}{c} \eta \\ \eta_1 \end{array} \right] (0,p\tau)$ η, η_1 integral characteristics.

To introduce the theorem from [Ko] we need the notion of a function order k of characteristics and characteristics $\left[\begin{array}{c} 0...0 \\ 0...0 \end{array} \right]$.

Definition 2.13. $f : C^g \times \mathbf{H_g} \mapsto C$ is an analytic function of order k and characteristics $\left[\begin{array}{c} 0...0 \\ 0...0 \end{array} \right]$ if the following relation is satisfied:

$$f\left(z+n+m^t\tau,\tau\right) = \exp\left\{2\pi i(-km^t z - km^t\tau m)\right\} f\left(z,\tau\right)$$

Let $k = p_1p_2$ and p_1, p_2 are even arbitrary numbers. We quote the following theorem from [Ko].

Theorem 2.14. *Let f be a function of characteristics $\left[\begin{array}{c} 0...0 \\ 0...0 \end{array} \right]$ and even order k. If $\left[\begin{array}{c} \mu \\ \mu' \end{array} \right]$ is an integral odd characteristics then the following identity is valid :*

$$\sum_{\nu,\nu',0\leq\nu_i\leq p_1,0\leq\nu_i'\leq p_2} (-1)^{\mu\nu-\mu'\nu'} f\left(\frac{\nu}{p_1}+\tau\frac{\nu'}{p_2}\right) = 0$$

and $\mu\nu = \sum_i \mu_i\nu_i$.

To obtain a modular equation for an odd number p, Choose $k = 2^{[log_2(p)]+1}, l = k-p$ and examine the function

$$f = \Theta \left[\begin{array}{c} 00...0 \\ 00...0 \end{array} \right]^l (z,\tau)\Theta \left[\begin{array}{c} 00...0 \\ 00...0 \end{array} \right] (pz,p\tau)$$

This function is of characteristics $\left[\begin{array}{c} 00...0 \\ 00...0 \end{array} \right]$ and order k. Define $p_1 = 2^{[log_2(p)]}, p_2 = 2$. Apply Theorem 2.14 to obtain the following:

Theorem 2.15. *(Modular equation for p) For any* $\left[\begin{array}{c}\mu\\ \mu'\end{array}\right]$ *odd integral characteristics:*

$$\sum_{\nu,\nu',0\leq\nu_i'\leq p_1,0\leq\nu_i\leq 1}(-1)^{\mu\nu-\mu'\nu'}\Theta\left[\begin{array}{c}\nu_i\\ \frac{2\nu_i'}{p_1}\end{array}\right]^l(0,\tau)\Theta\left[\begin{array}{c}\nu_i\\ \frac{2\nu_i'}{p_1}\end{array}\right](0,p\tau)=0$$

where $p_1=2^{[log_2(p)]}$. The main theorem of the paper is given below:

Theorem 2.16. *There exist explicit equation connecting*

$$\Theta\left[\begin{array}{c}\chi_i\\ \chi_i'\end{array}\right]^l(0,\tau),\ \textit{and}\ \Theta\left[\begin{array}{c}\chi_i\\ \chi_i'\end{array}\right]^l(0,p\tau)$$

χ_i,χ_i' *are g integral characteristics.*

Proof. According to Theorem 2.15:

$$\sum_{\nu,\nu',0\leq\nu_i'\leq p_1,0\leq\nu_i\leq 1}(-1)^{\mu\nu-\mu'\nu'}\Theta\left[\begin{array}{c}\nu_i\\ \frac{2\nu_i'}{p_1}\end{array}\right]^l(0,\tau)\Theta\left[\begin{array}{c}\nu_i\\ \frac{2\nu_i'}{p_1}\end{array}\right](0,p\tau)=0$$

and $\left[\begin{array}{c}\mu\\ \mu'\end{array}\right]$ is an odd characteristics.

Applying Theorem 2.9 we conclude that $\Theta\left[\begin{array}{c}\nu_i\\ \frac{2\nu_i'}{p_1}\end{array}\right]^l(0,\tau)$ is constructible from $\Theta\left[\begin{array}{c}\eta\\ \eta_1\end{array}\right](0,\tau)$ and η,η_1 is an integral characteristics. Replace

$$\Theta\left[\begin{array}{c}\nu_i\\ \frac{2\nu_i'}{p_1}\end{array}\right]^l(0,\tau),\Theta\left[\begin{array}{c}\nu_i\\ \frac{2\nu_i'}{p_1}\end{array}\right]^l(0,p\tau)$$

with the corresponding radical expression involving $\Theta\left[\begin{array}{c}\eta\\ \eta_1\end{array}\right](0,\tau)$ to conclude the result.

□

3. The one dimensional case

In this section we explain the connection between the theory developed in the last section and the usual one dimensional theory of modular equations.

Let E be an elliptic curve given in Legendre's normal form $y^2=x(x-1)(x-\lambda)$. if τ denotes the period that is induced by E, we have the following expression for λ as function of τ :

$$\lambda(\tau)=\frac{\Theta^4\left[\begin{array}{c}1\\ 0\end{array}\right](0,\tau)}{\Theta^4\left[\begin{array}{c}0\\ 0\end{array}\right](0,\tau)} \tag{8}$$

Recall the identity

$$\Theta^4 \begin{bmatrix} 0 \\ 0 \end{bmatrix} (0,\tau) = \Theta^4 \begin{bmatrix} 1 \\ 0 \end{bmatrix} (0,\tau) + \Theta^4 \begin{bmatrix} 0 \\ 1 \end{bmatrix} (0,\tau) \tag{9}$$

dividing both sides by $\Theta^4 \begin{bmatrix} 0 \\ 0 \end{bmatrix} (0,\tau)$ we see that

$$1 - \lambda(\tau) = \frac{\Theta^4 \begin{bmatrix} 0 \\ 1 \end{bmatrix} (0,\tau)}{\Theta^4 \begin{bmatrix} 0 \\ 0 \end{bmatrix} (0,\tau)} \tag{10}$$

Now set $\Theta^4 \begin{bmatrix} 0 \\ 0 \end{bmatrix} (0,\tau) = \theta_0(\tau), \Theta^4 \begin{bmatrix} 0 \\ 1 \end{bmatrix} (0,\tau) = \theta_1(\tau), \Theta^4 \begin{bmatrix} 1 \\ 0 \end{bmatrix} (0,\tau) = \theta_2(\tau)$
We write:

$$\sqrt[4]{\lambda(\tau)} = \frac{\theta_1(\tau)}{\theta_0(\tau)}, \tag{11}$$

and

$$\sqrt[4]{1 - \lambda(\tau)} = \frac{\theta_2(\tau)}{\theta_0(\tau)}. \tag{12}$$

The proof of the main theorem in the last section applied to the one dimensional case produces a homogenous radical expression of the form:

$$F\left(\theta_i(\tau)\theta_j(p\tau)\right) = 0. \tag{13}$$

Divide each term of the expression by $\theta_0(\tau)\theta_0(p\tau)$. Using the definition of $\lambda(\tau)$ replace its quotient by $\lambda(\tau)$ and $\lambda(p\tau)$ respectively. We obtain the following classical theorem:

Theorem 3.1. *There exists an algebraic relation between $\lambda(\tau)$ and $\lambda(p\tau)$.*

Note that the proof of this theorem we obtain does not use the usual modular group theory. The proof is constructive and provides an alternative way to construct modular equations for any p.

As an example consider the case $p = 3$ then applying the algorithm we obtain:

$$\theta_0(\tau)\theta_0(3\tau) = \theta_1(\tau)\theta_1(3\tau) + \theta_2(\tau)\theta_2(3\tau) \tag{14}$$

divide the two sides of the last equation by $\theta_0(\tau)\theta(3\tau)$ we obtain the classical modular equation:

$$1 = \sqrt[4]{\lambda(\tau)\lambda(3\tau)} + \sqrt[4]{(1-\lambda(\tau))\,(1-\lambda(3\tau))}.$$

More details can be found in [Bo].

4. Examples

We apply our theory to two cases as an example:

4.1. **p=3, g=2.** In this section we outline the result of the method applied above to $p = 3$ for $g = 2$. First we observe that for $g = 2$ we have 6 odd and 10 even characteristics. We conclude immediately that overall we have 6 modular equations for each p. Using the Theorem 2.7 we write for $p = 3$:

Theorem 4.1. *For any* $\begin{bmatrix} \mu \\ \mu' \end{bmatrix}$ *odd integral characteristics the following identities are true:*

$$\sum_{\nu,\nu',0\leq\nu_i'\leq 1,0\leq\nu_i\leq 1} (-1)^{\mu\nu-\mu'\nu'}\Theta\begin{bmatrix} \nu_i \\ \nu_i' \end{bmatrix}^l(0,\tau)\Theta\begin{bmatrix} \nu_i \\ \nu_i' \end{bmatrix}(0,3\tau) = 0$$

To achieve a more compact set of identities we rely on the classification of identities of power 4 achieved in [AK].

Definition 4.2. Let

$$0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, 1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, 2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, 3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Using the last definition we can write any 2 dimensional characteristics using the vectors above. For example:

$$03 = \begin{bmatrix} 01 \\ 01 \end{bmatrix}$$

We define matrix A the encodes even characteristics in the following way :

$$A = \begin{pmatrix} 11 & 01 & 10 \\ 22 & 20 & 02 \\ 33 & 21 & 22 \end{pmatrix}, \begin{pmatrix} 00 \end{pmatrix}$$

Then the following classification of 2 dimensional theta identities is given in [AK]

Theorem 4.3. *There are two types of relations for theta functions in power 4:*

- *Type I - corresponds to generalized diagonals of matrix A on the one side and the characteristics* $\begin{pmatrix} 00 \end{pmatrix}$ *on the other for example the following identity is true :*
$$\Theta_{00}^4 = \Theta_{11}^4 + \Theta_{20}^4 + \Theta_{12}^4(B)$$
- *Type II - Correspond to* 2×2 *sub matrices of A putting 2 not the same row and not the same column entries on each side of the equation. For example:*
$$\Theta_{11}^4 + \Theta_{20}^4 = \Theta_{01}^4 + \Theta_{22}^4$$

The identities of power 4 of theta constants were obtained applying the Gauss elimination procedure to the matrix that defines the power 4 identities of theta functions [AK]. Since the same matrix governs the identities involving terms of the form $\Theta(0,\tau)\Theta(0,3\tau)$ an immediate corollary of the last theorem is the following classification of identities involving prime $p = 3$:

Theorem 4.4. *There are two types of relations for theta functions involving the prime 3:*

- *Type I - corresponds to generalized diagonals of matrix A on the one side and the characteristics* $\begin{pmatrix} 00 \end{pmatrix}$ *for example the following identity is true :*

$$\Theta_{00}(0,\tau)\Theta_{00}(0,3\tau) = \Theta_{11}(0,\tau)\Theta_{11}(0,3\tau) + \Theta_{20}(0,\tau)\Theta_{20}(0,3\tau) + \Theta_{12}(0,\tau)\Theta_{12}(0,3\tau)$$

- *Type II - Correspond to 2×2 sub matrices of A putting 2 not the same row and not the same column entries on each side of the equation. For example:*

$$\Theta_{11}(0,\tau)\Theta_{11}(0,3\tau) + \Theta_{20}(0,\tau)\Theta_{20}(0,3\tau) = \Theta_{01}(0,\tau)\Theta_{01}(0,3\tau) + \Theta_{22}(0,\tau)\Theta_{22}(0,3\tau)$$

This gives an algebraic way to evaluate $\Theta_{ij}(0,3\tau)$ from $\Theta_{ij}(0,\tau)$. Compare with [CKL].

4.2. **p=7, g=2.** Let us write the equations explicitly for $p = 7$. According to the recipe outlined above we need to choose $k = 8$ and thus our basic function will be:

$$f = \Theta \begin{bmatrix} 00 \\ 00 \end{bmatrix}(z,\tau)\Theta \begin{bmatrix} 00 \\ 00 \end{bmatrix}(7z,7\tau)$$

for $p = 7$ Now we use theorem 2.14 to obtain the 6 equations for each odd characteristics. For example if the odd characteristics is

$$\begin{bmatrix} 10 \\ 10 \end{bmatrix}$$

The equation is:

$$\sum_{0\le\nu_1,\nu_2\le 1, 0\le\nu_1',\nu_2'\le 2} (-1)^{\nu_1-\nu_1'}\,\Theta \begin{bmatrix} \nu_1\nu_2 \\ \frac{\nu_1'}{2}\frac{\nu_2'}{2} \end{bmatrix}(0,\tau)\Theta \begin{bmatrix} \nu_1\nu_2 \\ \frac{\nu_1'}{2}\frac{\nu_2'}{2} \end{bmatrix}(0,7\tau) = 0$$

To translate the equation into equation involving integral characteristics we use the duplication formulas. For example we write:

$$\Theta^2 \begin{bmatrix} 11 \\ \frac{1}{2}\frac{1}{2} \end{bmatrix}(0,\tau) = \Theta \begin{bmatrix} 11 \\ 11 \end{bmatrix}(0,2\tau)\Theta \begin{bmatrix} 00 \\ 00 \end{bmatrix}(0,\tau) + \Theta \begin{bmatrix} 00 \\ 11 \end{bmatrix}(0,2\tau)\Theta \begin{bmatrix} 11 \\ 00 \end{bmatrix}(0,\tau)$$

Since the other two theta functions with integral characteristics are equal to 0. Similar formulas hold for the other functions of the form : $\Theta \begin{bmatrix} \nu_1\nu_2 \\ \frac{\nu_1'}{2}\frac{\nu_2'}{2} \end{bmatrix}(0,\tau)$ Substituting we obtain formulas involving integral characteristics at point $\tau, 2\tau, 14\tau$. To reduce to equations involving $\tau, 7\tau$ we apply the formulas from corollary 2.5 to functions

$$\Theta \begin{bmatrix} \nu_1\nu_2 \\ \nu_1'\nu_2' \end{bmatrix}(0,2\tau)$$

and ν_i, ν_i' are integral characteristics.

References

[AK] R.Adin, Y.Kopeliovich, Short Eigenvectors and Multidimensional Theta Functions, *Linear Algebra and Appl.* **257**(1)(1997) 49-63

[Bo] J. Borwein, P.Borwein, II *and the AGM*, A Wiley Interscience publication, 1987

[CKL] R.Carls, D.Kohel and D.Lubicz, Higher Dimensional 3-Adic CM Construction *Preprint*

[CL] R.Carls, and D.Lubicz, A p-adic quasi quadratic point counting algorthm http://arxiv.org/PS_cache/arxiv/pdf/0706/0706.0234v2.pdf

[FK1] H. Farkas, Y. Kopeliovich, New Theta Constant Identities *Israel Journal of Mathematics* **82**(1)(1993) 133-140

[FK2] H. Farkas, Y.Kopeliovich, New Theta Constant Identities II *Proceeding of AMS.***123**(4)(1995) 1009-1020

[GS] P.Gaudry, E.Schost, Modular Equations for Hyperelliptic curves http://www.csd.uwo.ca/~eschost/publications/papier2.pdf

[Ko] Y. Kopeliovich, Multi Dimensional Theta Constant Identities *Journal of Geometric Analysis* **8** (4)(1998) 571-581

[Ma] M.Madsen, A general framework for p - adic point counting and applications to elliptic curves on Legendre form http://www.imf.au.dk/publications/pp/2004/imf-pp-2004-2.pdf

[Me] J.-F.Mestre, Notes on Talk given at seminar of Cryptography at Rennes 2002. http://www.math.univ-rennes1.fr/crypto/2001-02/mestre.ps

[Mu] D. Mumford, *Tata Lectures on Theta II* (Progress in Mathematics, Birkhauser 1984)

[RF] H. Rauch and H.Farkas *Theta functions with application to Riemann Surfaces* (William and Wilkins Balt. Md. 1974)

Yaacov Kopeliovich
540 Madison Avenue, 6 -th floor
New York NY 10022
Email: ykopeliovich@yahoo.com,
ykopeliovich@medtolife.com

ALBANIAN JOURNAL
OF MATHEMATICS
Volume 1, Number 4, Pages 283–295
ISSN 1930-1235: (2007)

ON THE EXTREMAL REGULAR DIRECTED GRAPHS WITHOUT COMMUTATIVE DIAGRAMS AND THEIR APPLICATIONS IN CODING THEORY AND CRYPTOGRAPHY

V. A. USTIMENKO

Abstract. We use term regular directed graph (r. d. g.) for the graph of irreflexive binary relation with the constant number outputs (or inputs) for each vertex. The paper is devoted to studies of maximal size $E_R(d,v)$ of r. d. g. of order v without commutative diagrams formed by two directed passes of length $< d$ with the common starting and ending points. We introduce the upper bound for $E_R(d,v)$, which is one of the analogs of well known Even Circuit Theorem by P. Erdös'. The Erdös' theorem establish the upper bound on maximal size of simple graphs without cycles of length $2n$. It is known to be sharp for the cases $n = 2, 3 and 5$ only. The situation with the upper bound for $E_d(v)$ is different: we prove that it is sharp for each $d \geq 2$. We introduce the girth of directed graph and establish tight upper and lower bounds on the order of directed cages, i.e. directed regular graphs of given girth and minimal order. The studies of regular directed graphs of large size (or small order) without small commutative diagrams, especially algebraic explicit constructions of them, are motivated by their applications to the design of turbo codes in Coding Theory and cryptographical algorithms. We introduce several new algebraic constructions of directed extremal graphs based on biregular generalized polygons, family of directed graphs of large girth with fixed degree.

1. Introduction

According to Bourbaki the graph (or directed graph) is the pair V (vertex set) and subset Φ in the Cartesian product $V \times V$ (see [24] for more general definitions). We refer to element $v \in V$ as vertex (state in automata theory).

We use term arc (or arrow as in automata theory) for the element $(a,b) \in \Phi$. We refer to $(a,b) \in \Phi$ as arc (arrow) from a to b, Element a and b are starting and ending vertex of the arc (a,b). We say that (a,b) is output of vertex a and b is input of b. As it follows from above definition graph has no multiple arcs. The cardinalities of V and Φ are the order and size of the graph, respectively.

Graph is simple if Φ is symmetric and anti-reflexive relation. The information about simple graph can be given by edge i. e. set of kind $\{a,b\}$, where (a,b) is an arc. Graphically simple graph has no loops and multiple edges. In case of simple graph term size used for the number of edges within the graph.

The classical extremal graph theory studies extremal properties of simple graphs. Let F be family of graphs none of which is isomorphic to a subgraph of the graph Γ. In this case we say that Γ is F-free. Let P be certain graph theoretical property.

Key words and phrases. directed cages, directed graphs of large girth, directed small world graphs, bounds on order of directed cages, turbo codes, graph based cryptography.

By $\mathrm{ex}_P(v,F)$ we denote the greatest number of edges of F-free graph on v-vertices, which satisfies property P. If P is just a property to be simple graph we omit index P and write $\mathrm{ex}(v,F)$. The missing definitions in extremal graph theory the reader can find in [3].

This theory contains several important results on $\mathrm{ex}(v,F)$, where F is a finite collection of cycles of different length [3], [25]. The following statement had been formulated by P. Erdös'.

Let C_n denote the cycle of length n. Then

$$\mathrm{ex}(v,C_{2k}) \le Cv^{1+1/k} \tag{1.1}$$

where C is independent positive constant. For the proof of this result and its generalizations see [5], [8]. In [7] the upper bound

$$\mathrm{ex}(v,C_3,C_4,\dots,C_{2k},C_{2k+1}) \le (1/2)^{1+1/k}v^{1+1/k} + O(V) \tag{1.2}$$

had been established for all integers $k \ge 1$.

Both bounds are known to be sharp for $k = 2,3,5$ in other cases the question on the sharpness is open (see [3], [1] and further references).

The girth of the simple graph is the minimal length of its cycle. So the above bound is the restriction on the size of the graph on v vertices of girth $\ge n$. Graphs of high girth, i.e. graphs which size is close to the above upper bounds can be used in Networking and Operation Research (see [3]) and Cryptography.

The generalizations (or analogs) of classical extremal graph theory on directed graphs require certain restrictions on inputs or outputs of the graph. Really, the graph DK_v: $P \cup L = V$, $P \cap L = \emptyset$, $|V| = v$, $\Phi = P \times L$ of order $O(v^2)$ has no directed cycles or commutative diagrams.

In [38] the above results on maximal size of the graphs generalized on the case of balanced graphs, when for each vertex $a \in V$ cardinalities of $\mathrm{id}(v) = \{x \in V | (a,x) \in \phi\}$ and $\mathrm{od}(v) = \{x \in V | (x,a) \in \phi\}$ are same. We refer to numbers $\mathrm{id}(v)$ and $\mathrm{od}(v)$ as input degree and output degree of vertex v in the graph, respectively.

Let Γ be directed graph. The *pass* between vertices a and b is the sequence $a = x_0 \to x_1 \to \dots x_s = b$ of length s, where x_i, $i = 0,1,\dots,s$ are distinct vertices. We refer to he minimal s among all passes between a and b as output distance $\mathrm{odist}(a,b)$. we assume $\mathrm{odist}(a,b) = \infty$ in case of absence of passes from a to b.

We say that the pair of passes $a = x_0 \to x_1 \to \dots \to x_s = b$, $s \ge 1$ and $a = y_0 \to y_1 \to \dots \to y_t = b$, $t \ge 1$ form an (s,t)- commutative diagram $O_{s,t}$ if $x_i \ne y_j$ for $0 < i < s$, $0 < j < t$. Without loss of generality we assume $s \ge t$ and refer to the number s as the rank of $O_{s,t}$. The directed cycle with s arrows we denote as $O'_{s,0}$. We will count directed cycles as commutative diagram.

The minimal parameter $s = \max(s,t)$ of the commutative diagram $O_{s,t}$ with $s + t \ge 3$ in the binary relation graph Γ we call the *girth indicator* of the Γ and denote it as $\mathrm{gi}(\Gamma)$. It can be infinity as in case of DK_v.

Notice that directed graph does not contain diagrams $O_{1,1}$, because there are no multiple edges.

We assume that the *girth* $g(\Gamma)$ of directed graph Γ with the girth indicator $d+1$ is $2d+1$ if it contains commutative diagram $O_{d+1,d}$. If there are no such diagrams we assume that $g(\Gamma)$ is $2d+2$.

In the case of symmetric irreflexive relations it agrees with the standard definition of the girth of simple graph i.e the length of its minimal cycle.

Let F be a list of directed graphs and P be some graph theoretical property. By $\mathrm{Ex}_P(v,F)$ we denote the greatest number of arrows of F-free directed graph on v vertices satisfying to property P (graph without subgraphs isomorphic to graph from F).

Let $\mathrm{E}_P = \mathrm{E}_P(d,v) = \mathrm{Ex}_P(v, O_{s,t}, s+t \geq 3 | 2 \leq s \leq d)$ be the maximal size (number of arrows) of the balanced binary relation graphs with the girth indicator $> d$.

The main result of [38] is the following statement. If B be the property to be the balanced directed graph, then

$$v^{1+1/d} - O(v) \leq \mathrm{E}_B(d,v) \leq v^{1+1/d} + O(v) \tag{1.3}$$

Notice, that the size of symmetric irreflexive relation is the double of the size of corresponding simple graph. because undirected edge of the simple graph corresponds to two arrows (arcs) of $O_{2,0}$.

If P is the property to be a graph of symmetric irreflexive relation then

$$\mathrm{Ex}_P(v, O_{s,t}, s+t \geq 3 | 2 \leq s \leq d) = 2\mathrm{ex}(v, C_3, \ldots, C_{2d-1}, C_{2d}),$$

because undirected edge of the simple graph corresponds to two arrows of $O_{2,0}$. So equality (1, 3) implies the following inequality

$$\mathrm{ex}(v, C_3, C_4, \ldots, C_{2k}) \leq (1/2)v^{1+1/k} + O(V) \tag{1.4}$$

we evaluate the maximal size of the directed graph of order v with the girth indicator $> d$ which does not contain commutative diagrams $O_{d+1,d}$, as well. The inequality (1.2) is the corollary from such evaluation.

We can see that studies of extremal properties of balanced graphs with the high girth indicator and studies of $ex(v, C_3, \ldots, C_n)$ are far from being equivalent. Really, the sharpness of the Erdös' bound (1.1) and bounds (1.2) and (1.4) up to magnitude for $k = 8$ and $k \geq 12$ are old open questions (see [1], [3]) .

The regularity R of graph (V, Φ) means that either for each vertex $a \in V$ sets $\{x | (v,x) \in \Phi\}$ are same or for each $a \in V$ set $\{x | (x,v) \in \Phi\}$ are same. We will prove that substitution of property R instead of B leads to correct inequality:

$$v^{1+1/d} - O(v) \leq \mathrm{E}_R(d,v) \leq v^{1+1/d} + O(v) \tag{1.5}$$

The family of directed graphs G_i, $i = 1, \ldots$ with average output degree k_i and order is the family of large girth if the girth indicator of G_i is $\geq \log_{k_i}(v)$. It agrees well with the standard definition for the simple graphs. In case of balanced or regular graphs of large girth their size is close to the upper bounds (1. 3) and (1. 5). The following directions of applied data security are motivations of studies of extremal properties of regular or balanced graphs of large girth.

1.1. **LDPS and Turbo Codes and graphs of large girth.** Low-density parity-check (LDPC) codes were originally introduced in his doctoral thesis by Gallager in 1961 [11]. Since the discovery of Turbo codes in 1993 by Berrou, Glavieux, and Thitimajshima [4], and the rediscovery of LDPC codes by Mackay and Neal in 1995 [21], there has been renewed interest in Turbo codes and LDPC codes, because their error rate performance approaches asymptotically the Shannon limit. Much research is devoted to characterizing the performance of LDPC codes and designing codes that have good performance. Commonly, a graph, the Tanner graph (see [26] and further references), is associated with the code and an important parameter affecting the performance of the code is the girth of its Tanner graph. In [23], [12],

[13] authors consider the design of structured regular LDPC codes whose Tanner graphs have large girth. The regularity and structure of LDPC codes utilize memory more efficiently and simplify the implementation of LDPC coders. The Tanner graph is a special type of graph, a bipartite graph, where the nodes divide into two disjoint classes with edges only between nodes in the two different classes. The impotence of the studies of undirected regular bipartite graphs with large girth for the design of turbo codes is discussed in [23].

Large girth speeds the convergence of iterative decoding and improves the performance of LDPC codes, at least in the high SNR range, by slowing down the on setting of the error floor. Large size of such graphs implies fast convergence.

1.2. **Cryptography.** The cryptographical properties of infinite families of simple graphs of large girth with the special coloring of vertices during the last 10 years (see [31],[34], [33], [35] and further references). Such families can be used for the development of cryptographical algorithms (on symmetric or public key modes). Only few families of simple graphs of large unbounded girth and arbitrarily large degree are known.

Paper [35], [38] is devoted to the more general theory of directed graphs of large girth and their cryptographical applications. It contains new explicit algebraic constructions of infinite families of such graphs. It is shown that they can be used for the implementation of secure and very fast symmetric encryption algorithms. The symbolic computations technique allow us to create a public key mode for the encryption scheme based on algebraic graphs. The information on the implementations if such algorithms can be found in [30],[34], [15], [29] (case of simple graphs) and [35], [37], [16]. Last two papers compare speed of the graph based algorithms with the speed of RC4 and DES.

2. On the upper bounds for size of the regular graphs the with high girth indicator

Let Γ be the graph of irreflexive binary relation Φ on the vertex set V and the following property R holds:

for each vertex $v \in V$ the input degrees $\mathrm{id}(v) = |\{x|(x,v) \in \Phi\}| = k$ or $\mathrm{od}(v) = |\{x|(v,x) \in \Phi\}| = k$ for some positive number $k \geq 2$.

As it follows from property B for balanced graph Φ the cardinality $\{(x,y,z)|(x,y) \in \Phi and (y,z) \in \Phi\}$ is $D = \sum_{v \in V} ({k_v}^2)$. So the number of random walk with two arrows from random vertex v is D/v. Any random walk in this graph can be viewed as the branching process with $\sqrt{(}D/v)$ branches from each node.

The bound $\mathrm{E}_B(d,v) \leq v^{1+1/d} + O(v)$ is based on the studies of such branching process corresponding to the passes of length $\leq d$ of the rooted tree. The definitions of such branching process, expectation operator and the confidence interval the reader can find in the book [14] by Karlin and Taylor.

In our case of regular graphs we can use straightforward combinatorial counting.

Theorem 1.

$$\mathrm{E}_R(d,v) \leq v^{1+1/d} + O(v) \tag{2.1}$$

$$\mathrm{Ex}_R(v, O_{d+1,d}, O_{s,t} | 3 \leq s \leq d) \leq (1/2)^{1/d} v^{1+1/d} + O(v) \tag{2.2}$$

Proof. Let $\Gamma = \Gamma_i$, $i = 1, \dots, v$ be the family of regular graphs corresponding to irreflexive binary relations $\phi = \phi_i$ with the girth indicator i which is $> d$ and maximal possible number of edges on $v = v_i$ vertices. Without loss of generality we may assume $od(v) = k$ for each vertex of Γ. Let us chose the vertex x_0 and consider the totality V_r of all vertices from Γ, such that $r = \mathrm{odist}(x_o, v) \leq d$. We can use the branching process in counting of $v_r = |V_r|$: graph has no loops, so v_1 is k. One link of the vertex $x \in V_2$ may correspond to diagram $O_{2,0}$ so $v_2 \geq k(k-1)$. Induction on i we are getting $v_i = k(K-1)^{i-1}$ for $i = 1, \dots, d$.

We have $(k-1)^d \leq k(k-1)^d \leq v_d \leq v$. Thus, $(k-1)^d \leq v$, $(k-1) \leq v^{1/d}$ $E(\Gamma) = v \times (k \leq v \times (v^{1/d} + 1) = v^{1+1/d} + v$. So we proved (2.1).

Let us assume now that graphs Γ_i, $i = 1, \dots$ do not contain commutative diagrams $O_{d+1,d}$. Let us consider the arc $v_1 \to v_2$ in the graph Γ and two rooted trees T_1 and T_2 with roots v_1 and v_2, respectively. Let P_i be the sets of vertices at the distance d and from the vertex v_i, $i = 1, 2$ The absence of commutative diagrams listed in the theorem insure that $|P_1 \cap P_2| = 0$ and $|P_1 \cup P_2| = (k-1)^d$. Thus $2(q-1)^d \leq v$. So for the size of the graph $E(\Gamma)$ is $\leq v \times ((v/2)^d + 1) = (1/2)^d v^d + v$. □

3. On the sharpness of the bound

The diameter is the minimal length d of the shortest directed pass $a = x_0 \to x_1 \to x_2 \cdots \to x_d$ between two vertices a and b of the directed graph. We will say that graph is k-regular, if each vertex of G has exactly k outputs. Let F be the infinite family of k_i regular graphs G_i of order v_i and diameter d_i. We say, that F is a family of small world graphs if $d_i \leq C\log_{k_i}(v_i)$, $i = 1, \dots$ for some independent on i constant C. The definition of small world graphs and related explicit constructions the reader can find in [3]. For the studies of simple small world graphs without small cycles see [10].

Let M be a finite set, $m = |M| \geq 2$. We define M_k, $m \geq k+2$ as the totality of tuples $(x_1, x_2, \dots x_k) \in M^k$, such that $x_i \neq x_j$ for each pair $(i, j) \in M^2$. Let us consider the binary relation $\phi = \phi_k(m)$ on M_k consisting of all pairs of tuples $((x_1, \dots, x_m)\ , (y_1, \dots, y_m))$, such that $y_i = x_{i+1}$ for $i = 1, \dots, k-1$ and $y_m \neq x_i$ for each $i \in \{1, \dots, k\}$. The corresponding directed graph $\Gamma = \Gamma_i(m)$ has order $m(m-1)\dots(m-k+1)$, each vertex has $m-k$ input and output arrows.

Proposition 2. *The girth indicator and diameter of the graph $\Gamma_k(m)$ is $k+1$ and $2k$, respectively. The girth of the graph is $2d+1$*

Proof. Let us consider the $O_{s,t}$, $0 \leq t \leq s \leq k$, $s \geq 1$ of the graph $\Gamma_k(m)$ with the starting point a $= (a_1, a_2, \dots, a_k)$. Let $\mathrm{a}_x = (a_2, a_3, \dots, a_k, x)$ be the neighbor of $\tilde{a}$ within the pass P_x of the diagram of length s. Notice that x is different from a_i, $i = 1, 2, \dots, k$. Let P be other pass of the diagram. If length t of P is zero, we assume that P consist of one vertex a. The first component of ending point w of the P_x equals to x. But the first component of each vertex for each vertex of the pass P is either element of $\{a_1, a_2, \dots, a_k\}$ (case $t < s$ or element y, $y \neq x$ (case $t = s$). But w has to be the vertex of P as well. So we are getting a contradiction. Thus, we proved that the girth indicator of the graph is $> k$.

Notice that w $= (x, x_1, \dots, x_{k-1})$, where $x \neq a_i$, $i = 1, \dots, k$, $x_i \neq a_j$, $J = i+1, i+2, \dots, k$, $j = 1, \dots, k-1$. We can add vertex $(x_1, x_2, \dots, x_{k_1}, x_k)$, consider

the following specialization of variables $x_i = a_i$ for $i = 1, 2, \dots, k$ and obtain the diagram $O'_{0,k+1}$. So the girth indicator of the graph is $k+1$.

Let us consider the pass of length $2k$ starting from a and going throw w and $(x_1, x_2, \dots, x_k)$ as above. It contains the following tuples:

$$(x_2, \dots, x_k, y_1),\ (x_3, \dots, x_k, y_1, y_2), \dots (x_k, y_1, \dots, y_{k-1}).$$

The only requirement on distinct elements $X_k, Y_1, \dots, y_k$ is x_k *nex* and x can be arbitrarily element from the complement of $\{a_1, \dots, a_k\}$. If $m \geq k+2$, then arbitrary point of M_k can be reached from a via the pass as above and diameter of the graph is bounded by $2k$. It is clear that there is no pass of length $2k-1$ between a and element of kind $(Z_1, \dots, z_{k-1}, a_k)$. So $\mathrm{diam}(\Gamma_k(m)) = 2m$.

□

Corollary 3. *Let F be the family of graphs $\Gamma_m(k)$, $m = k+2, k+3, \dots$. Then it is a family of directed small world graphs, the size of the members of this family is on the bound (2.1) of theorem 1.*

Really, $\Gamma_m(k)$ has degree $m-k$, order $v = m(m-1)\dots(m-k+1)$. So $\log_{m-k}(v)$ is some constant $> k$. So diameter of graphs from the family is bounded by $2\log_{m-k}(v)$. The size of $\Gamma_m(k)$ is $v(m-k)$. We have $(m)^k \geq V$. So $E(\Gamma_m(k)) \geq v[(v^{1/k}) - k] = v^{1+1/k} - kv$.

Let us consider the bipartite analog $\Gamma' = \Gamma'_k(m)$ of the graph $\Gamma = \Gamma_k(m)$ Let M be a finite set, $m = |M| \geq 2$. Let P (point set) and L (line set) are two copies of the vertex set M_k, $m \geq k+2$ of the graph Γ. We will use the brackets and parenthesis for the tuples from P and L, respectively.

Let$\Gamma' = \Gamma'_k(m)$ be the graph of binary relation on $P \cup L$ consisting of all pairs of tuples $((x_1, \dots, x_m), [y_1, \dots, y_m)$ or $(x_1, \dots, x_m, (y_1, \dots, y_m))$, such that $y_i = x_{i+1}$ for $i = 1, \dots, k-1$ and $y_m \neq x_i$ for each $i \in \{1, \dots, k\}$. The corresponding directed graph $\Gamma' = \Gamma'_k(m)$ has order $2m(m-1)\dots(m-k+1)$, each vertex has $m-k$ input and output arrows.

Proposition 4. *The girth indicator and diameter of the graph $\Gamma'_k(m)$ is $k+1$ and $2k+1$, respectively. The graph does not contain commutative diagram $O_{k+1,k}$.The girth of the graph is $2d+2$.*

Proof. The graph does not contain $O_{k+1,k}$ because of the ending point of the diagram can not be point and line at same time.The evaluation of the girth indicator and diameter can be done similarly to the evaluation in the proof of proposition 1. □

Corollary 5. *Let F' be the family of graphs $\Gamma'_m(k)$, $m = k+2, k+3, \dots$. Then it is a family of directed small world graphs, the size of the members of this family is on the bound (2.2) of theorem 1.*

Really, $\Gamma'_m(k)$ has degree $m-k$ and order $v = 2m(m-1)\dots(m-k+1)$. We have $(m-k)^K \leq m(m-1)\dots(m-k+1)$. So $k \leq \log_{m-k}(m(m-1)\dots(m-k+1))$. Thus $2k+1 < 3k \leq 3\log_{m-k}(m(m-1)\dots(m-k+1)) < 3\log 2m(m-1)\dots(m-k+1) = 3\log_{m-k}(v)$

The size of $\Gamma'_m(k)$ is $v(m-k)$. We have $(m)^k \geq m(m-1)\dots(m-k+1) = v/2$. So $m > (1/2)^k v^{1/k}$. Thus $E(\Gamma'_m(k)) \geq v((1/2)^{1/k} v^{1/k}) - k] = (1/2)^{1/k} v^{1+1/k} - kv$.

4. On the directed graphs without commutative diagrams of rank $< d$ of minimal order

Recall that (k,g)-cage is a simple graph of degree k, girth g of minimal order $v(k,g)$. The following objects are analogies of classical cages.

Definition 6. *We refer to the directed graph with the girth g, output degree k and minimal order $u(k,g)$ as directed (k,g)-cage.*

As it follows from the definition of directed (k,g)-cage

Theorem 7. *The following hold:*

$$(k+d)(k+d-1)\dots k \geq u(2k+1,d) \geq 1+k(k-1)+\dots k(k-1)^{d-1},$$

$$2[(k+d)(k+d-1)\dots k] \geq u(2k+2,d) \geq (1+(k-1)+\dots (k-1)^d)+(k-1)^d$$

Proof. Let Γ be directed graph with k-outputs for each vertex and girth indicator d, then the branching process Branch starting with the chosen vertex a gives $s = 1+k+k^(k-1)+\dots k(k-1)^d$ different vertices. So we prove (i).

Let b satisfies to $a \to b$. We can consider $K-1$ output arcs (a,x) from a, which are different from (a,b). The branching process starting from each element x b gives at least $(K-1)+\dots (k-1)^{d-1}$ passes of length $\leq d-1$. This way we are getting set T of elements of distance $(d-1)$ from a. Let us consider arcs of kind (b,y), $y \neq a$. The branching process from y gives us $(q-1)+(q-1)^{d-1}$ at distance $d-2$ from y. Together with b we have $1+(q-1)+\dots (q-1)^{d-1}$ elements at distance $\leq d-1$ from b. This set has empty intersection with T because of absence of commutative diagrams $O_{d+1,d}$. So we have at least $(1+(k-1)+\dots (k-1)^d)+(k-1)^d$ different vertices of the graph.

□

Proposition 8. *Let Γ be directed cage with the output degree ≥ 3 of order v and girth indicator d.*

(i) If its girth is $2d+1$, then the size E of the graph satisfies to the following inequality

$$v^{1+1/d} - kv \leq E \leq v^{1+1/d} + v$$

(ii) if its girth is $2d+2$, then the size E of the graph satisfies to the following inequality

$$(1/2)^{1/d}v^{1+1/d} - kv \leq E \leq (1/2)^{1/d}v^{1+1/d} + v$$

Let P be some property of directed regular graphs and $u_P(k,g)$ be the minimal order of graph with the output degree K and the girth indicator g. It is clear that $U_P(k,g) \geq U(k,g)$. So $v(m,g) \geq u(m,g)$, in particular. The following statement follows immediately from the above inequalities.

Corollary 9. *Let s be the property to be simple graph. Then*

a) $v(k,2d+1) = u_S(k,2d+1) \geq u(k,2d+1) \geq 1+k+k(k-1)+\dots+k(k-1)^{d-1}$,

b) $v(k,2d+2) = u_s(k,2d+2) \geq U(k,2d+2) \geq (1+(k-1)+\dots+(k-1)^d)+(k-1)^d$

The above lower bound for $g = 2d+2$ can be improved by Tutte inequality $v(k,2d+2 \geq 2(1+(k-1)+\dots(k-1)^d)$ (see [BCN]). The Tutte's lover bound for $v(k,2d+2)$ is same with (b). The upper and lower bound for $U(k,g)$ are quite tight, both of them are given by polynomial expression in variable k of kind $k^d + f(k)$, where $d = [(q-1)/2]$ and $\deg f(x) \leq d-1$. The situation with the known upper

bound on the order of cages is different, such bound is quite far from the lower one (see [18]).

Cages of odd girth with the order on the Tutte's bound are known as Moore graphs. There are only finite examples of Moore graphs. Well known finite generalized m-gons are examples of cages of even girth (see next section of the paper).

From the existence of the k-regular Moore graph of girth $2d+1$ $(2d+2)$ follows $U(k,d) = v(k,2d+1) = 1 + k(k-1) + \dots k(k-1)^{d-1}$ $(u(k,d) = v(k,2d+1) = 2(1 + (k-1) + \dots (k-1)^d)$, respectively.

There is a finite number of Moore graphs of order v of odd girth. Some infinite families of Moore graphs of even girth are known (see [6] or next section).

Proposition 10. *Let A be the property to be the graph of antisymmetric relation Φ i.e. $(a,b) \in \Phi$ implies that $(b.a)$ is not in Φ. Then*
(i) $(k+d)(k+d-1)\dots k \geq u_A(2k+1,d) \geq 1 + k + k^2 + \dots k^{d-1}$,
(ii) $2[(k+d)(k+d-1)\dots k] \geq u_A(2k+2,d) \geq [1+k+k^2+\dots k^{d-1}]+(k-1)k^{d-1}$.

The bounds (i) and (ii) are valid for balanced antisymmetric regular graphs.

5. Algebraic explicit constructions of extremal regular directed graphs with the fixed girth indicator

We shall use term of *algebraic graph* for the of graph $\Gamma(K)$ of binary relation Φ, such that the vertex set $V(\Gamma) = V(K)$ is an algebraic variety over commutative ring K of dimension ≥ 1 and for each vertex $v \in V$ the neighborhoods $\mathrm{In}(v) = \{x|(x,v) \in V\}$ and $\mathrm{Ou}(v) = \{x|(v,x) \in V\}$ are algebraic varieties over K of dimension ≥ 1 as well (see [2] for the case of simple graphs).

We shall use term *the family of directed graphs of large girth* for the family of regular graphs Γ_i with output degree k_i and order v_i such that their girth indicator $d_i = \mathrm{gi}(\Gamma_i)$ are $\geq c \log_{k_i}(v_i)$, where c is the independent on i constant. So the size of such graphs is quite close to the bound (2.1) or (2.2).

We say that Γ_i form a family of asymptotical directed cages of odd (even) girth if $v_i = ki^{d_i} + o(ki^{d_i})$ ($v_i = 2ki^{d_i} + o(ki^{d_i})$. It is clear that asymptotical cages of even or odd girth are families of graphs of large girth.

In this section we consider examples of families of algebraic graphs of large girth with fixed girth indicator, asymptotical directed cages of odd and even girth, in particular.

E. Moore [11] used term *tactical configuration* of order (s,t) for biregular bipartite simple graphs with bidegrees $s+1$ and $r+1$. It corresponds to incidence structure with the point set P, line set L and symmetric incidence relation I. Its size can be computed as $|P|(s+1)$ or $|L|(t+1)$.

Let $F = \{(p,l)|p \in P, l \in L, pIl\}$ be the totality of flags for the tactical configuration with partition sets P (point set) and L (line set) and incidence relation I. We define the following irreflexive binary relation ϕ on the set F:

$((l_1,p_1),(l_2,p_2)) \in \phi$ if and only if $p_1 I l_2$, $p_1 \neq p_2$ and $l_1 \neq l_2$. Let $F(I)$ be the binary relation graph corresponding to ϕ. The order of $F(I)$ is $|P|(s+1)$ (or $|L|(t+1)$ We refer to it as *directed flag graph* of I.

Lemma 11. *Let (P,L,I) be a tactical configuration with bidegrees $s+1$ and $t+1$ of girth $g \geq 4k$. Then the girth indicator of directed graph $F(I)$ with the output an input degree st is $> k$.*

Let (P, L, I) be the incidence structure corresponding to regular tactical configuration of order t.

Let $F_1 = \{(l, p) | l \in L, p \in P, lIp\}$ and $F_2 = \{[l, p] | l \in L, p \in P, lIp\}$ be two copies of the totality of flags for (P, L, I). Brackets and parenthesis allow us to distinguish elements from F_1 and F_2. Let $DF(I)$ be the directed graph (double directed flag graph) on the disjoint union of F_1 with F_2 defined by following rules:

$(l_1, p_1) \to [l_2, p_2]$ if and only if $p_1 = p_2$ and $l_1 \neq l_2$,
$[l_2, p_2] \to (l_1, p_1)$ if and only if $l_1 = l_2$ and $p_1 \neq p_2$.

Lemma 12. *Let (P, L, I) be a regular tactical configuration of degrees s of girth $g \geq 2m$. Then the girth indicator of double directed graph $DF(I)$ with the output an input degree s is $> m$.*

Generalized m-gons $GP_m(r, s)$ of order (r, s) were defined by J. Tits in 1959 (see [15], [16] and survey [14]) as tactical configurations of order (s, t) of girth $2m$ and diameter m. According to well known Feit - Higman theorem a finite generalized m-gon of order (s, t) has $m \in \{3, 4, 6, 8, 12\}$ unless $s = t = 1$.

The known examples of generalized m-gons of bidegrees ≥ 3 and $m \in \{3, 4, 6, 8\}$ include rank 2 incidence graphs of finite simple groups of Lie type (see [4]). The regular incidence structures are $I_{1,1}(3, q)$ for $m = 3$ (groups $A_2(q)$), $I_{1,1}(4, q)$, $m = 4$ (groups $B_2(q)$) and $I_{1,1}(6, q)$, $m = 6$ (group $G_2(q)$). In all such cases $s = t = q$, where q is prime power.

The biregular but not regular generalized m-gons have parameters $s = q^\alpha$, $t = q^\beta$, where q is a prime power. The list is below: relation $I_{2,1}(4, q)$, $s = q^2, t = q$, q is arbitrary large prime power for $m = 4$; $I_{3,2}(6, q)$, $s = q^3, t = q^2$, where $q = 3^{2k+1}$, $k > 1$ for $m = 6$; $I_{2,1}(8, q)$, $s = q^2, t = q$, $q = 2^{2k+1}$ for $m = 8$. For each triple of parameters (m, s, t) listed above there is an edge transitive generalized m-gon of order (s, t) related to certain finite rank 2 simple group of Lie type. in case of $m = 3$ (projective planes. in particular) and $m + 4$ (generalized quadragons) some infinite families of graphs without edge transitive automorphism group are known. The following two lemmas can be obtained immediately from the axioms of generalized polygon.

Lemma 13. *Let (P, L, I) be the generalized $2k$-gon of order (r, s). Then*
$|P| = \sum_{t=0,k-1} (r^t s^t + r^{t+1} s^t)$, $|L| = \sum_{t=0,k-1} (s^t r^t + s^{t+1} r^s)$.

Lemma 14. *Let (P, L, I) be regular generalized m-gon of degree $q+1$. Then $|P| = |L| = 1 + q + \cdots + q^{m-1}$.*

Corollary 15. *For each $m = 3, 4, 6$ and prime p the family $F_m(q)$, $q = p^n$, $n = 1, \ldots$ of edge transitive polygons is an algebraic family over F_p of cades of girth $2m$ of degree $q + 1$ with the order on the Tutte's lower bound.*

Let (P, L, I) be generalized m-gon of order (s, t), $s \geq 2$, $t \geq 2$ and $e = \{(p, l\}$, $(p \in P, l \in L, pIl)$ be chosen edge of this simple graph.

Let $\mathrm{S}_e = \mathrm{Sch}_e(I)$ be the restriction of incidence relation I onto $P' \cup L'$ where P' (L') is the totality of points (lines) at maximal distance from p (l, respectively). It can be shown that (P', L', S_e) is a tactical configuration of degree $(s - 1, t - 1)$. Let us refer to (P', L', S_e) as Schubert graph. If the generalized polygon is edge-transitive its Schubert graph is unique up to isomorphism. In this case Schubert

graph corresponds to the restriction of incidence relation onto the union of two "largest large Schubert cells", i. e. orbits of standard Borel subgroups of highest dimension.

Proposition 16. *For each $S_m(p)$ $m = 3, 4, 6$ and prime p the family of Schubert graphs $S_m(p)$ of regular generalized m-gons $F_m(q)$ is algebraic over F_p family of asymptotical cages of even girth with the order $2q^{m-1}$ and degree q.*

The extremal properties of finite generalized polygons, their Schubert graphs and some of their induced subgraphs have been considered in [32].

Remark. The girth of $S_m(q)$ is $2m$ for "sufficiently large" parameter q.

Let (P, L, I) be a regular tactical configuration of order (t, t). The *double configuration* $I' = DT(I)$ is the incidence graph of the following incidence structure $(P', L', I') : P' = F(I) = \{(p, l)|p \in P, l \in L, pIl\}$, $L' = P \cup L$, $f = (p, l)Ix$, $x \in L'$ if $p = x$ or $l = x$. It is clear that the order of tactical configuration I' is $(1, t)$. If (P, L, I) is a generalized m-gon, then (P', L', I') is a generalized $2m$-gon.

Proposition 17. *(i) If the girth of regular tactical configuration (P, L, I) of degree $s + 1$ is $2t$, then the girth of $DT(I)$ is $4t$. The order of $DT(I)$ is $(s, 1)$.*

(ii) Let (P, L, I) be regular generalized m-gon, then $DT(I)$ is generalized $2m$-gom.

Corollary 18. *The configurations $DT(I) = I^2(m, q)$ for known regular m-gons, $m = 3, 4, 6$ of degree $q + 1$, q is a prime power, are generalized $2m$-gons of order $(1, q)$.*

Theorem 19. *(i) Let $I_{s.t}(m, q)$, $m \geq 4$ be the incidence relation of one of the known edge transitive m-gons defined over the field F_q, $q = p^n$, p is prime number.*

Then for each tuple (m, s, t, p) the family of directed flag-graphs $F^n(m, s, t, p)$ for generalized m-gon of order $(q^s, q^t$ is an algebraic over F_p family of asymptotic cages of odd girth. The girth indicator of each graph from the family is $m/2 + 1$ and the girth is $m + 1$ $(5, 7, 9)$.

(ii) Let $S_{s.t}(m, q)$, $m \geq 4$ be the Schuberst graph of the incidence relation $I_{s.t}(m, q)$ of one of the known edge transitive m-gons defined over the field F_q, $q = p^n$, p is prime number.

Then for each tuple (m, s, t, p) the family of directed flag-graphs $SF^n(m, s, t, p)$ for $S_{s.t}(m, q)$ is an algebraic over F_p family of asymptotic cages of odd girth. The girth indicator of each graph from the family is $m/2 + 1$ and the girth is $m + 1$.

(iii) Let $I_{1.1}(m, q)$ be the incidence relation of one of the known edge transitive regular m-gons defined over the field F_q, $q = p^n$, p is prime number. Then for each pair (m, p) the family $DF(m, p)$ of double flag graphs $DF(m, i) = DF(I_{1.1}(m, p^i))$, $i = 1, \ldots$ is an algebraic over F_p family of directed asymptotic cages of even girth. The girth indicator of each $DF(m, s)$ is $m + 1$ and the girth is $2m + 2$ $(8, 10.14)$.

(4i) Let $I^2(m, q)$ be the incidence relation of double tactical configuration of regular generalized m gon defined over F_q, $q = p^n$, p is prime. Then for each pair (m, p) the family $F(m, p)$ of directed flag-graphs $F^n(m, p)$, $n = 1, \ldots$ is an algebraic over F_p family of directed graphs of large girth. The girth indicator of each graph is $m + 1$ and girth is $2m + 1$ $(7, 9, 13)$.

Regular finite generalized polygons have been used in works of R. Tanner on Coding Theory. The applications of biregular generalized polygons and their Schubert graphs to Cryptography the reader can find in [33]. Paper [37] devoted to

cryptographical algorithms based on nonsymmetric directed asymptotical cages as above.

6. On the constructions of families of nonsymmetric directed graphs of large girth with fixed degree

The concept of family of simple graphs of large girth of fixed degree had been introduced by P. Erdös' in the late 50th.

The first explicit examples of families of simple graphs with large girth of arbitrary large degree were given by Margulis. The constructions were Cayley graphs $X^{p,q}$ of group $SL_2(Z_q)$ with respect to special sets of $q+1$ generators, p and q are primes congruent to 1 mod4. The family of $X^{p,q}$ is not a family of algebraic graphs because the neighborhood of each vertex is not an algebraic variety over F_q. For each p, graphs $X^{p,q}$, where q is running via appropriate primes, form a family of small world graph of unbounded diameter (see [19],[20]).

The fist family of connected algebraic simple graphs over F_q of large girth and arbitrarily large degree had been constructed in [17]. These graphs $CD(k,q)$, k is an integer ≥ 2 and q is odd prime power had been constructed as connected component of graphs $D(k,q)$ defined earlier. For each q graphs $CD(k,q)$, $k \geq 2$ form a family of large girth with $\gamma = 4/3\log_{q-1}q$.

Two new examples of families of simple algebraic graphs of large girth and arbitrary large degree the reader can find in [36]. Papers [34], [30], [15], [29] devoted to software packages of cryptographical algorithms based on simple graphs.

For each commutative ring the infinite family of directed graphs of large girth with fixed degree has been constructed in [34]. The cryptographical algorithms based on such directed graphs the reader can find in [34], [36]. Paper [15] devoted to implementations of graph based fast stream ciphers corresponding to antisymmetric relations.

References

[1] C.T. Benson, *Minimal regular graphs of girth eight and twelve*, Canadien Journal of Mathematics, (18):1091- 1094, 1966.

[2] N. Biggs, *Algebraic Graph Theory* (2nd ed), Cambridge, University Press, 1993.

[3] B. Bollobás, *Extremal Graph Theory*, Academic Press, London, 1978.

[4] C. Berrou, A. Glavieux and P. Thitimajshima, *Near Shannon limit errorcorrectingcoding and decoding: turbocodes*, ICC 1993, Geneva, Switzerland, pp. 10641070, May 1993.

[5] J.A. Bondy and M.Simonovits, *Cycles of even length in graphs*, J. Combin.Theory, Ser. B, 16 (1974) 87-105.

[6] A. Brower, A. Cohen, A. Nuemaier, *Distance regular graphs*, Springer, Berlin, 1989.

[7] P. Erdös', M. Simonovits, *Compactness results in extremal graph theory*, Combinatorica 2 (3), 1982, 275-288.

[8] W. Faudree, M. Simonovits, *On a class of degenerate extremal graph problems*, Combinatorica 3 (1), 1983, 83-93.

[9] W. Feit, D. Higman *The nonexistence of certain generalised polygons*, J. of Algebra 1 (1964), 114-131.

[10] V. Futorny, V. Ustimenko, *On Small World Semiplanes with Generalised Schubert Cells*, Acta Applicandae Mathematicae, N4, 2007 (already available online).

[11] R. G. Gallager, Lowdensity paritycheck codes, IRE Transactions on Information Theory, vol. IT8, pp. 2128, Jan. 1962.

[12] P. Guinand and J. Lodge, "Tanner Type Codes Arising from Large Girth Graphs", Proceedings of the 1997 Canadian Workshop on Information Theory (CWIT '97), Toronto, Ontario, Canada, pp. 5-7, June 3-6, 1997.

[13] P. Guinand and J. Lodge, *Graph Theoretic Construction of Generalized Product Codes*, Proceedings of the 1997 IEEE International Symposium on Information Theory (ISIT '97), Ulm, Germany, p. 111, June 29-July 4, 1997.

[14] S. Karlin, H.M. Taylor, *A first course in stochastic processes*, Academic Press, New York, 1975.

[15] Yu Khmelevsky, V Ustimenko, Practical aspects of the Informational Systems reengineering, The South Pacific Journal of Natural Science, volume 21, 2003, p.75-21 (together with Yu. Khmelevsky), www.usp.ac.fj/spjns/volume21

[16] J. Kotorowich, V. Ustimenko, *On the implementation of cryptoalgorithms based on algebraic graphs over some commutative rings*, Proceedings of the international conferences "Infinite particle systems, Complex systems theory and its application, Kazimerz Dolny, Poland, 2005-2006 (to appear).

[17] F. Lazebnik, V. A. Ustimenko and A. J. Woldar, *A New Series of Dense Graphs of High Girth*, Bull (New Series) of AMS, v.32, N1, (1995), 73-79.

[18] (with F. Lazebnik and A. Woldar) New upper bound on the order of cages, Electronic Journal of Combinatorics, Volume 4 (1997), No. 2, Paper R13.

[19] A. Lubotsky, R. Philips, P. Sarnak, *Ramanujan graphs*, J. Comb. Theory., 115, N 2., (1989), 62-89.

[20] G. A. Margulis, *Explicit construction of graphs without short cycles and low density codes*, Combinatorica, 2, (1982), 71-78.

[21] D. J. C. MacKay and R. N. Neal, *Good Codes based on very sparse matrices*, In "Cryptography and Coding", 5th IMA Conference, Lecture Notes in Computer Science, v. 1025, 1995, pp. 110-111.

[22] E. H. Moore, *Tactical Memoranda*, Amer. J. Math., v.18, 1886, 264-303.

[23] Jose M. F. Moura, Jin Lu, and Haotian Zhang, *Structured LDPC Codes with Large Girth*, IEEE Signal Processing Magazine, vol. 21:1, pp.42-55, January 2004. Included in Special Issue on Iterative Signal Processing for Communications

[24] R. Ore, *Graph Theory*, Wiley, London, 1971.

[25] M. Simonovits *Extremal Graph Theory*, Selected Topics in Graph Theory 2 (L.W. Beineke and R.J. Wilson, eds), Academic Press, London, 1983, 161-200.

[26] R. Michiel Tanner, *A recursive approach to low density codes*, IEEE Trans. on Info Th., IT, 27(5):533-547, Sept.1984.

[27] J. A. Thas, *Generalised polygons*, in F. Buekenhout (ed), Handbook in Incidence Geometry, Ch. 9, North Holland, Amsterdam, 1995.

[28] J. Tits, *Sur la trialite et certains groupes qui s'en deduisent*, Publ. Math. I.H.E.S, 2 (1959), 15-20.

[29] A. Touzene, V. Ustimenko, *Graph Based Private Key Crypto System*, International Journal on Computer Research, Nova Science Publisher, volume 13 (2006), issue 4, 12p.

[30] V. Ustimenko, D. Sharma, *CRYPTIM: system to encrypt text and image data*, Proceedings of International ICSC Congress on Intelligent Systems 2000, Wollongong, 2001, 11pp.

[31] V. A. Ustimenko, *Coordinatisation of regular tree and its quotients*, in "Voronoi's impact on modern science, eds P. Engel and H. Syta, book 2, National Acad. of Sci, Institute of Matematics, 1998, 228p.

[32] V. Ustimenko, A. Woldar, *Extremal properties of regular and affine generalised polygons as tactical configurations*, 2003, European Journal of Combinatorics, 24, 99-111.

[33] V. Ustimenko, *Graphs with Special Arcs and Cryptography*, Acta Applicandae Mathematicae, 2002, vol. 74, N2, 117-153.

[34] V. Ustimenko, *CRYPTIM: Graphs as tools for symmetric encryption*, In Lecture Notes in Comput. Sci., 2227, Springer, New York, 2001.

[35] V. Ustimenko, *On the extremal graph theory for directed graphs and its cryptographical applications*, In: T. Shaska, W. C. Huffman, D. Joener and V. Ustimenko, Advances in Coding Theory and Cryptography, Series on Coding Theory and Cryptology, vol. 3, 181-200 (2007).

[36] V. A. Ustimenko, *Linguistic Dynamical Systems, Graphs of Large Girth and Cryptography*, Journal of Mathematical Sciences, Springer, vol. 140, N3 (2007), pp 412-434.

[37] V. Ustimenko *On the graph based cryptography and symbolic computations*, Serdica journal of computing, N1, 2007 (to appear).

[38] V. Ustimenko *On the extremal balanced binary relation graphs of high girth*, Proceedings of the international conferences "Infinite particle systems", Complex systems theory and its application, Kazimerz Dolny, Poland, 2005-2006 (to appear).

Vasyl Ustimenko
University of Maria Curie-Sklodowska
Lublin, Poland
Email: vasyl@golem.umcs.lublin.pl

ALBANIAN JOURNAL
OF MATHEMATICS
Volume 1, Number 4, Pages 297–319
ISSN 1930-1235: (2007)

SOME OPEN PROBLEMS IN COMPUTATIONAL ALGEBRAIC GEOMETRY

TANUSH SHASKA

To my wonderful children
Rachel, Adrianna, Eva, and Besianna.

Abstract. The development of computational techniques in the last decade has made possible to attack some classical problems of algebraic geometry from a computational viewpoint. In this survey, we briefly describe some open problems of computational algebraic geometry which can be approached from such viewpoint. Some of the problems we discuss are the decomposition of Jacobians of genus two curves, automorphisms groups of algebraic curves and the corresponding loci in the moduli space of algebraic curves $\mathcal{M}_g$, inclusions among such loci, decomposition of Jacobians of algebraic curves with automorphisms, invariants of binary forms and the hyperelliptic moduli, theta functions of curves with automorphisms, etc. We decompose Jacobians of genus 3 curves with automorphisms and determine the inclusions among the loci for algebraic curves with automorphisms of genus 3 and 4.

Contents

Key words and phrases. computational algebraic geometry, algebraic curves, automorphisms, Hurwitz spaces.

1. Introduction

Computational algebraic geometry is a very active and rapidly growing field, with many applications to other areas of mathematics, computer science, and engineering. In this survey, we will focus on some open problems of algebraic geometry which can be approached by a computational viewpoint.

The first version of this paper appeared in 2003 in the ACM, *SIGSAM Bulletin*, *Comm. Comp. Alg.*, see [26]. It was a list of problems on algebraic curves. Some of those problems were solved and many papers were written based on that modest paper. Since then I have updated the list with new problems and have included problems on higher dimensional varieties.

In the first part, we focus on algebraic curves and revisit some of the problems of the 2003 list. We report on some progress made on some of the problems and work done in other problems. Most notably, there are many papers generated on the field of moduli versus the field of definition problem.

This survey is organized in two parts. In Part 1 we survey some open problems, related to algebraic curves, which can be attacked using computational techniques. In Part 2 we discuss a couple of problems about higher dimensional varieties. When we say computational techniques we don't necessarily mean only Groebener basis and elimination techniques. Instead our understanding of computational geometry includes computational group theory, numerical methods using homotopy techniques, complex integration, combinatorial methods, etc.

Part 1 contains sections 2-7. In the second section we describe genus 2 curves with split Jacobians. There are many papers written on these topic going back to Legendre and Jacobi in the context of elliptic integrals. The problem we suggest is to compute the moduli space of covers of degree 5, 7 from a genus 2 curve to an elliptic curve. This problem is completely computational and could lead to some better understanding of some conjectures on elliptic curves; see Frey [9].

In section three, we discuss the automorphism groups of algebraic curves. There has been some important progress on this topic lately, however much more can be done. Extending some of the results to positive characteristic would be important. Further we suggest computing the equations of Hurwitz curves of genus 14 and 17.

In section 4 we study automorphism groups of algebraic curves defined over fields of positive characteristics. Determining such groups has theoretical applications as well as applications in coding theory, cryptography, etc. While a complete answer to this problem might still be out of reach, it seems that it is possible to determine such groups and equations of curves for special classes of curves.

In section 5 we study the decomposition of Jacobians of curves via the automorphisms of curves. We completely deal with the case $g = 3$. As far as we are aware this is the first time this result appears in the literature. Since now we have full

knowledge of the list of automorphism groups for any genus it seems possible (and reasonable) that such decomposition be determined for reasonably small genus (say $g \leq 10$).

In section 6, we go back to the problem of invariants of binary forms. The reader might find interesting the fact that we believe that the result of Shioda is not correct.

Denote by $S(n, r)$ the graded ring of invariants of homogeneous polynomials of order r in n variables over C. Shioda determines the structure of $S(2, 8)$, which turns out to be generated by nine invariants $J_2, \cdots, J_{10}$ satisfying five relations; see [35, On the graded ring of invariants of binary octavics. Amer. J. Math. 89 1967 1022–1046.]. He also computes explicitly five independent syzygies, and determines the corresponding syzygy-sequence. We believe such relations are not correct and should be determined using computational algebra tools.

In section 7 we introduce the problem of determining relations among theta functions for algebraic curves with automorphisms. This is a long project of the author and his collaborators; see for example [22] for more details. There is some progress on this topic lately, by some attempts to generalize Thomae's formula for cyclic curves; see [22] for references. Such formula, when known, expresses branch points of the cover $\mathcal{X} \to \mathbb{P}^1$ in terms of theta functions. Since such branch points for cyclic curves are easily determined then it becomes a problem of computational algebra to determine such relations.

Part 2 contains two sections. The problems stated here are of particular interest to the author. No effort is made to have a comprehensive list of problems in higher dimensional varieties. The first problem is to determine the degree of a rational map and the second problem is that of parameterizing surfaces.

Most of the problems suggested in this survey and software programs connected to them can be found in [8]. We have tried to make available most of the computer files where the computations are performed at:

http://algcurves.albmath.org/

Throughout this paper it is assumed that the reader is familiar with computer algebra packages as GAP [10] and the library of small groups in GAP.

Notation: Throughout this paper an "algebraic curve" means the isomorphism class of the curve defined over an algebraically closed field, unless otherwise stated.

Part 1. Algebraic curves

Algebraic curves are one of the oldest and most studied branches of algebraic geometry and indeed one of the most studied areas of mathematics. They provide some of the most exciting problems of classical mathematics. Meanwhile, they have found applications in many different areas of science and technology, such as computer vision, coding theory, cryptography, quantum information, biomathematics, etc. However, amazingly enough, there are some vary basic questions about algebraic curves of deep theoretical interest which still elude the community of experts in this area of research. We briefly describe a few problems of interest which can be studied using computational techniques.

2. Genus 2 curves with (n,n)-split Jacobian

In this section we focus on genus 2 curves whose Jacobians are isogenous to a product of elliptic curves. These curves have been studied extensively in the 19-th century in the context of elliptic integrals by Legendre, Jacobi, Clebsch, Hermite, Goursat, Brioschi, and Bolza et al. In the late 20th century Frey and Kani, Kuhn, Gaudry and Schost, Shaska and Voelklein, and many others have studied these curves further. They are of interest for the arithmetic of genus 2 curves as well as elliptic curves. For a complete survey on this topic see [32, 3] where [32] focuses on the computational aspects and [3] on connections of such coverings to the elliptic integral, mathematical physics, etc.

Let C be a curve of genus 2 and $\psi_1 : C \longrightarrow E$ a map of degree n, from C to an elliptic curve E, both curves defined over $\mathbb{C}$. The degree n cover $\phi : C \to E$ induces a degree n cover $\phi : \mathbb{P}^1 \to \mathbb{P}^1$ such that the following diagram commutes

$$\begin{array}{ccc} C & \xrightarrow{\pi_C} & \mathbb{P}^1 \\ {\scriptstyle\psi}\downarrow & & \downarrow{\scriptstyle\phi} \\ E & \xrightarrow{\pi_E} & \mathbb{P}^1 \end{array}$$

Figure 1. The basic setup

Here $\pi_C : C \to \mathbb{P}^1$ and $\pi_E : E \to \mathbb{P}^1$ are the natural degree 2 covers. Let r be the number of branch points of the cover $\phi : \mathbb{P}^1 \to \mathbb{P}^1$. Then $r = 4$ or $r = 5$, with $r = 5$ being the generic case and $r = 4$ occurring for a certain 1-dimensional sub-locus of $\mathcal{L}_n$. We refer to the case $r = 5$ (resp., $r = 4$) as the **non-degenerate case**, resp., the **degenerate case**; see [25, Thm 3.1].

If $\psi_1 : C \longrightarrow E_1$ is maximal[1] (i.e., does not factor non-trivially) then there exists a maximal map $\psi_2 : C \longrightarrow E_2$, of degree n, to some elliptic curve E_2 such that there is an isogeny of degree n^2 from the Jacobian J_C to $E_1 \times E_2$. We say that J_C is (n,n)-decomposable. If the degree n is odd the pair (ψ_2, E_2) is canonically determined; see [25] for details.

We denote the moduli space of such degree n coverings $\phi : \mathbb{P}^1 \to \mathbb{P}^1$ by $\mathcal{L}_n$. It can be viewed also as the Hurwitz space of covers $\phi : \mathbb{P}^1 \to \mathbb{P}^1$ with ramification determined as in [25]. For our purposes, $\mathcal{L}_n$ will simply be the locus of genus 2 curves whose Jacobian is (n,n)-isogenous to a product of two elliptic curves.

The case $n = 2$ is a special case since the coverings $\phi : \mathbb{P}^1 \to \mathbb{P}^1$ are Galois. The locus $\mathcal{L}_2$ of these genus 2 curves is a 2-dimensional subvariety of the moduli space $\mathcal{M}_2$ and is studied in detail in [34]. An equation for $\mathcal{L}_2$ is already in the work of Clebsch and Bolza. In [34] we found a birational parametrization of $\mathcal{L}_2$ by affine 2-space to study the relation between the j-invariants of the degree 2 elliptic subfields. We found a 1-dimensional family of genus 2 curves having exactly two isomorphic elliptic subfields of degree 2; this family is parameterized by the j-invariant of these subfields.

[1]Some authors would call such a map **minimal covering.**

2.1. **Covers of odd degree.** If $n > 2$, the surface $\mathcal{L}_n$ is less understood. The case $n = 3$ was initially studied by Kuhn [15] where some computations for $n = 3$ were performed. The computation of the equation of $\mathcal{L}_3$ was a major computational effort. Computational algebra techniques (i.e., Groebner basis, Buchberger algorithm) and computational algebra packages (i.e, Maple, GAP) were used. Let $K = \mathbb{C}(C)$ be genus 2 function fields of C. The ellipic subcovers E_1, E_2 correspond to degree 3 elliptic subfields of K. The number of $Aut(K)$-classes of such subfields of fixed K is 0, 1, 2 or 4; see [29] for details. Also, an equation for the locus of such C in the moduli space of genus 2 curves is computed. It was the first time to explicitly compute such spaces and the results were obtained with the help of computer algebra.

The case $n = 5$ is studied in detail in [19]. This extends earlier work for $n = 2, 3$ in Shaska [25], [29], and Shaska/Völklein [34]. The cover $\phi : \mathbb{P}^1 \to \mathbb{P}^1$ has one of the ramification structures given in Table 1. This data lists the ramification indices > 1 over the branch points. E.g., in the last case there is one branch point that has exactly one ramified point over it, of index 3, and each of the other 3 branch points has exactly two ramified points over it, of index 2. The reader should check [25] for ramification structures of arbitrary degree.

non-degenerate:		$\big((2)^2, (2)^2, (2)^2, (2), (2)\big)$
degenerate:		
	I)	$\big((2)^2, (2)^2, (4), (2)\big)$
	II)	$\big((2)^2, (2)^2, (2)\cdot(3), (2)\big)$
	III)	$\big((2)^2, (2)^2, (2)^2, (3)\big)$

TABLE 1. ramification structure of ϕ

The main feature that distinguishes the case $n = 5$ from all other values $n > 5$ is that the cover ϕ does not determine $\psi : C \to E$ uniquely, but there is essentially **two** choices of ψ for a given ϕ. These two choices correspond to the two branch points of ϕ of ramification structure (2) (notation as in Table 1) – anyone of these two branch points can be chosen to ramify in E while the other doesn't. This phenomenon implies that the function field of $\mathcal{L}_5$ is a quadratic extension of the function field of the Hurwitz space parameterizing the covers ϕ.

The spaces $\mathcal{L}_n$ were studied by many authors in different contexts. The new results obtained in [29, 19] were the result of applying successfully computational tools and computer algebra. Continuing on the work of the above papers, we suggest the following problem:

Problem 1. Determine the locus $\mathcal{L}_n$ in $\mathcal{M}_2$ for $n = 7$. Further, determine the relation between the elliptic curves E_1 and E_2 in each case.

Using techniques from [34, 29] this becomes simply a computational problem. However, determining such loci requires the use of a Groebner basis algorithm. Computationally this seems to be difficult for $n = 7$. Notice that $n = 7$ is the first generic case of the problem since all degenerate cases occur.

2.2. **Covers of even degree.** The case when n is even is less studied. In this case there are several possible ramifications that can occur for the covering $\phi : \mathbb{P}^1 \to \mathbb{P}^1$; see [25] for the following theorem.

Theorem 2. *If n is an even number then the generic case for $\psi : C \longrightarrow E$ induce the following three cases for $\phi : \mathbb{P}^1 \longrightarrow \mathbb{P}^1$:*

I.: $\Big((2)^{\frac{n-2}{2}}, (2)^{\frac{n-2}{2}}, (2)^{\frac{n-2}{2}}, (2)^{\frac{n}{2}}, (2)\Big)$

II.: $\Big((2)^{\frac{n-4}{2}}, (2)^{\frac{n-2}{2}}, (2)^{\frac{n}{2}}, (2)^{\frac{n}{2}}, (2)\Big)$

III.: $\Big((2)^{\frac{n-6}{2}}, (2)^{\frac{n}{2}}, (2)^{\frac{n}{2}}, (2)^{\frac{n}{2}}, (2)\Big)$

Each of the above cases has the following degenerations (two of the branch points collapse to one)

I.: (1) $\Big((2)^{\frac{n}{2}}, (2)^{\frac{n-2}{2}}, (2)^{\frac{n-2}{2}}, (2)^{\frac{n}{2}}\Big)$

(2) $\Big((2)^{\frac{n-2}{2}}, (2)^{\frac{n-2}{2}}, (4)(2)^{\frac{n-6}{2}}, (2)^{\frac{n}{2}}\Big)$

(3) $\Big((2)^{\frac{n-2}{2}}, (2)^{\frac{n-2}{2}}, (2)^{\frac{n-2}{2}}, (4)(2)^{\frac{n-4}{2}}\Big)$

(4) $\Big((3)(2)^{\frac{n-4}{2}}, (2)^{\frac{n-2}{2}}, (2)^{\frac{n-2}{2}}, (2)^{\frac{n}{2}}\Big)$

II.: (1) $\Big((2)^{\frac{n-2}{2}}, (2)^{\frac{n-2}{2}}, (2)^{\frac{n}{2}}, (2)^{\frac{n}{2}}\Big)$

(2) $\Big((2)^{\frac{n-4}{2}}, (2)^{\frac{n}{2}}, (2)^{\frac{n}{2}}, (2)^{\frac{n}{2}}\Big)$

(3) $\Big((4)(2)^{\frac{n-8}{2}}, (2)^{\frac{n-2}{2}}, (2)^{\frac{n}{2}}, (2)^{\frac{n}{2}}\Big)$

(4) $\Big((2)^{\frac{n-4}{2}}, (4)(2)^{\frac{n-6}{2}}, (2)^{\frac{n}{2}}, (2)^{\frac{n}{2}}\Big)$

(5) $\Big((2)^{\frac{n-4}{2}}, (2)^{\frac{n-2}{2}}, (2)^{\frac{n-4}{2}}, (2)^{\frac{n}{2}}\Big)$

(6) $\Big((3)(2)^{\frac{n-6}{2}}, (2)^{\frac{n-2}{2}}, (4)(2)^{\frac{n}{2}}, (2)^{\frac{n}{2}}\Big)$

(7) $\Big((2)^{\frac{n-4}{2}}, (3)(2)^{\frac{n-4}{2}}, (2)^{\frac{n}{2}}, (2)^{\frac{n}{2}}\Big)$

III.: (1) $\Big((2)^{\frac{n-4}{2}}, (2)^{\frac{n}{2}}, (2)^{\frac{n}{2}}, (4)(2)^{\frac{n}{2}}\Big)$

(2) $\Big((2)^{\frac{n-6}{2}}, (4)(2)^{\frac{n-4}{2}}, (2)^{\frac{n}{2}}, (2)^{\frac{n}{2}}\Big)$

(3) $\Big((2)^{\frac{n}{2}}, (2)^{\frac{n}{2}}, (2)^{\frac{n}{2}}, (4)(2)^{\frac{n-10}{2}}\Big)$

(4) $\Big((3)(2)^{\frac{n-8}{2}}, (2)^{\frac{n}{2}}, (2)^{\frac{n}{2}}, (2)^{\frac{n}{2}}\Big)$

The following problem is a natural extension of the techniques used in the odd degree case to the even degree.

Problem 3. Determine the loci $\mathcal{L}_n$ in $\mathcal{M}_2$ for $n = 4, 6$. Further, determine the relation between the elliptic curves E_1 and E_2 in each case.

We expect that the computation of such loci computationally to be easier than in the cases $n = 5$. The ramification structure in the case $n = 4$ is:

non-degenerate:		$\big((2)^2, (2), (2), (2), (2)\big)$
degenerate:		
	i)	$\big((2)^2, (2), (2), (2)^2\big)$
	ii)	$\big((2), (2), (2), (4), \ \big)$

TABLE 2. Ramification structures of ϕ for $n = 4$

Note that when n is even the choice of E_2, on contrary to the odd case, is not canonical. The reader who would like a more detailed survey on this topic should check [32, 3].

3. The automorphism group of algebraic curves

Computation of automorphism groups of compact Riemann surfaces is a classical problem that goes back to Schwartz, Hurwitz, Klein, Wiman and many others. Hurwitz showed that the order of the automorphism group of a compact Riemann surface of genus g is at most $84(g-1)$, which is known as the Hurwitz bound. Klein was mostly interested with the real counterpart of the problem, hence the term "compact Klein surfaces". Wiman studied automorphism groups of hyperelliptic curves and orders of single automorphisms.

The 20th century produced a huge amount of literature on the subject. Baily [6] gave an analytical proof of a theorem of Hurwitz: if $g \geq 2$, there exists a curve of genus g with non-trivial automorphisms. In other papers was treated the number of automorphisms of a Riemann surface; see Accola [2], Maclachlan [16], [17] among others. Accola [1] gives a formula relating the genus of a Riemann surface with the subgroups of the automorphism group; known as Accola's theorem. Harvey studied cyclic groups and Lehner and Newman maximal groups that occur as automorphism groups of Riemann surfaces.

A group of automorphisms of a compact Riemann surface X of genus g can be faithfully represented via its action on the Abelian differentials on X as a subgroup of $GL(g, \mathbf{C})$. There were many efforts to classify the subgroups G of $GL(g, \mathbf{C})$ that so arise, via the cyclic subgroups of G and conditions on the matrix elements of G. In a series of papers, I. Kuribayashi, A. Kuribayashi, and Kimura compute the lists of subgroups which arise this way for $g = 3, 4$, and 5.

By covering space theory, a finite group G acts (faithfully) on a genus g curve if and only if it has a genus g generating system. Using this purely group-theoretic condition, Breuer [7] classified all groups that act on a curve of genus ≤ 48. This was a major computational effort using the computer algebra system GAP. It greatly improved on several papers dealing with small genus, by various authors.

Of course, for each group in Breuer's list, all subgroups are also in the list. This raises the question how to pick out those groups that occur as the **full automorphism group** of a genus g curve.

Let G be a finite group, and $g \geq 2$. In [18] is studied the locus of genus g curves that admit a G-action of given type, and inclusions between such loci. We use this to study the locus of genus g curves with prescribed automorphism group G. We completely classify these loci for $g = 3$ (including equations for the corresponding curves), and for $g \leq 10$ we classify those loci corresponding to "large" G. Furthermore, such work has been continued by K. Magaard who has given complete answers for the list of groups (in characteristic 0) of algebraic curves of any genus g.

3.1. **Inclusion among the loci of curves with prescribed automorphism group.** Let H and G be groups which occur as automorphism groups of genus g algebraic curves such that $H < G$. If the cover $\mathcal{X} \to \mathcal{X}^H$ is obtained as a degeneration (collapsing of branch points) of the cover $\mathcal{X} \to \mathcal{X}^G$ then the locus $\mathcal{M}(g, H)$ (locus of curves in $\mathcal{M}_g$ with automorphism group H) is a sublocus of

$\mathcal{M}(g, G)$. We are avoiding signatures here, assuming that the reader is aware of the details of moduli spaces of covers and Hurwitz spaces; for details see [18].

Problem 4. Determine the inclusion among the loci $\mathcal{M}(g, G)$ for algebraic curves defined over $\mathbb{C}$ and genus $g \leq 10$.

The case of genus 2 is well known and can be found in [34] among others. Here we briefly describe the cases $g = 3, 4$.

3.1.1. *Genus 3.* Automorphism groups of genus 3 algebraic curves are well known. Furthermore, the subvarieties of $\mathcal{M}_3$ determined by group actions and inclusions among such loci are studied in detail; see [28, 27, 30, 12, 5] among others. There are 21 groups that occur as automorphism groups of genus 3 curves. Only two groups occur with two different signatures, namely $\mathbb{Z}_2$ and V_4. Such signatures distinguish between the hyperelliptic and non-hyperelliptic case. Hence, overall we have 23 cases, twelve of which belong to the non-hyperelliptic curves and the other eleven belong to the hyperelliptic locus.

Throughout this section we use the GAP identity of the library of small groups to identify the groups. We display the list of groups in Table 1 and Table 2. The equation of the family of curves, the signature, the number of involutions N_i and the number of conjugacy classes of involutions N_c are also displayed.

	Aut $(\mathcal{X}_g)$	$\overline{\text{Aut}}(\mathcal{X}_g)$	δ	equation $y^2 = f(x)$	Id.
1	$\mathbb{Z}_2$	$\{1\}$	5	$x(x-1)(x^5 + ax^4 + bx^3 + cx^2 + dx + e)$	$(2, 1)$
2	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$\mathbb{Z}_2$	3	$x^8 + a_3x^6 + a_2x^4 + a_1x^2 + 1$	$(4, 2)$
3	$\mathbb{Z}_4$	$\mathbb{Z}_2$	2	$x(x^2 - 1)(x^4 + ax^2 + b)$	$(4, 1)$
4	$\mathbb{Z}_{14}$	$\mathbb{Z}_7$	0	$x^7 - 1$	$(14, 2)$
5	$\mathbb{Z}_2^3$	D_4	2	$(x^4 + ax^2 + 1)(x^4 + bx^2 + 1)$	$(8, 5)$
6	$\mathbb{Z}_2 \times D_8$	D_8	1	$x^8 + ax^4 + 1$	$(16, 11)$
7	$\mathbb{Z}_2 \times \mathbb{Z}_4$	D_4	1	$(x^4 - 1)(x^4 + ax^2 + 1)$	$(8, 2)$
8	D_{12}	D_6	1	$x(x^6 + ax^3 + 1)$	$(12, 4)$
9	U_6	D_{12}	0	$x(x^6 - 1)$	$(24, 5)$
10	V_8	D_{16}	0	$x^8 - 1$	$(32, 9)$
11	$\mathbb{Z}_2 \times S_4$	S_4	0	$x^8 + 14x^2 + 1$	$(48, 48)$

TABLE 3. Aut (X_3) for hyperelliptic X_3

Hyperelliptic curves: Let $\mathcal{X}_g$ be a genus g hyperelliptic curve, $G :=$ Aut $(\mathcal{X}_g)$ the automorphism group, and $\sigma \in G$ its hyperelliptic involution. Then σ is in the center of G. The group $\overline{\text{Aut}}(\mathcal{X}_g) := G/\langle\sigma\rangle$ is called the *reduced automorphism group* of $\mathcal{X}_g$. It is a finite group of $PGL(2, \mathbb{C})$. Thus, $\overline{\text{Aut}}(\mathcal{X}_g)$ is isomorphic to a cyclic group, dihedral group, S_4, A_4, or A_5. Then, Aut $(\mathcal{X})g)$ is a degree 2 central extension of $\overline{\text{Aut}}(\mathcal{X}_g)$. Using these facts, for each $g \geq 2$ one determines the list of automorphism groups that occur, their signatures, and the parametric equation of corresponding curve. Moreover the inclusion among the loci $\mathcal{H}(G, C)$ is also known; see [27, 28, 11, 31, 11, 33, 24, 12] for details. Below we display the list of automorphism groups, the reduced automorphism group, a parametric equation of the curve, and the dimension of the corresponding locus. For the signature in each case

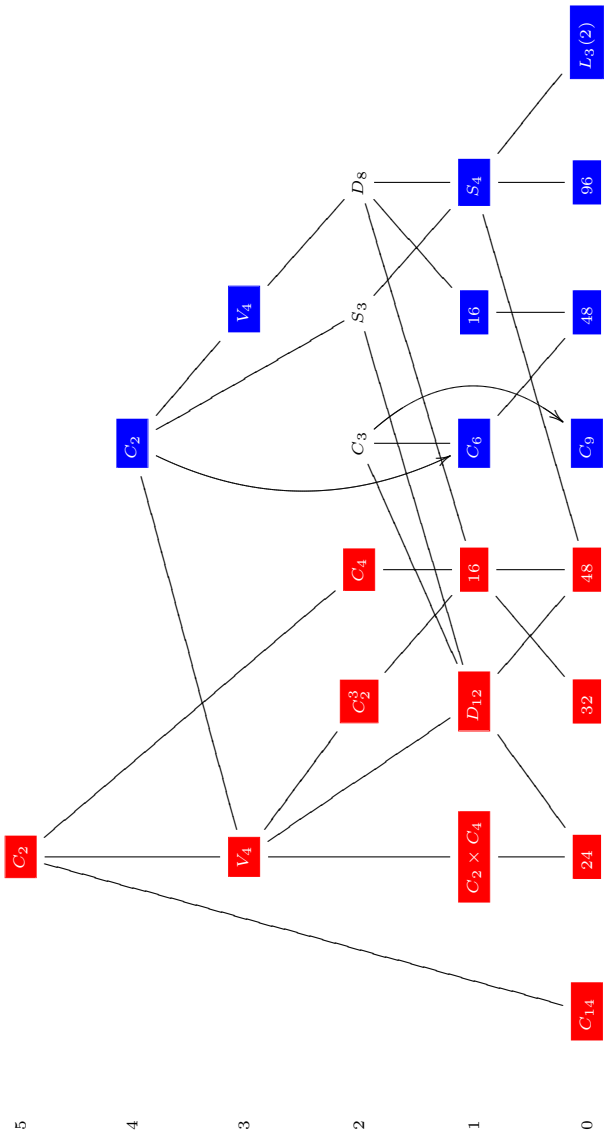
and other details the reader can check [12]. The subvarieties of $\mathcal{H}_3$ corresponding to each locus in the table are studied in details in [12]. In this paper we skip the details for the hyperelliptic moduli.

Non-hyperelliptic curves: A genus 3 non-hyperelliptic curve $\mathcal{X}$ is a ternary quartic. The group of automorphisms Aut $(\mathcal{X})$ is a finite group of order ≤ 168 with notably the Klein curve having automorphism group the simple group of order 168. The list of groups that occur as full automorphism groups of genus 3 curves are given in the table below. Each group is identified also with the Gap identity number. This number uniquely (up to isomorphism) determines the group in the library of small groups in GAP; see [10].

#	G	sig.	genus g_0	dim. δ	Id	N_i	N_c
1	V_4	(2^6)	0	3	(4,2)	3	3
2	D_8	(2^5)	0	2	(8,3)	5	3
3	S_4	$(2^3, 3)$	0	1	(24,12)	9	2
4	$C_4^2 \rtimes S_3$	$(2, 3, 8)$	0	0	(96,64)	15	2
5	16	$(2^3, 4)$	0	1	(16,13)	7	4
6	48	$(2, 3, 12)$	0	0	(48,33)	7	2
7	C_3	(3^5)	0	2	(3,1)	0	0
8	C_6	$(2, 3, 3, 6)$	0	1	(6,2)	1	1
9	C_9	$(3, 9, 9)$	0	0	(9,1)	0	0
10	$L_3(2)$	$(2, 3, 7)$	0	0	(168,42)	21	1
11	S_3	$(2^4, 3)$	0	2	(6,1)	3	1
12	C_2	(2^4)	1	4	(2,1)	1	1

TABLE 4. Automorphism groups of genus 3 non-hyperelliptic curves

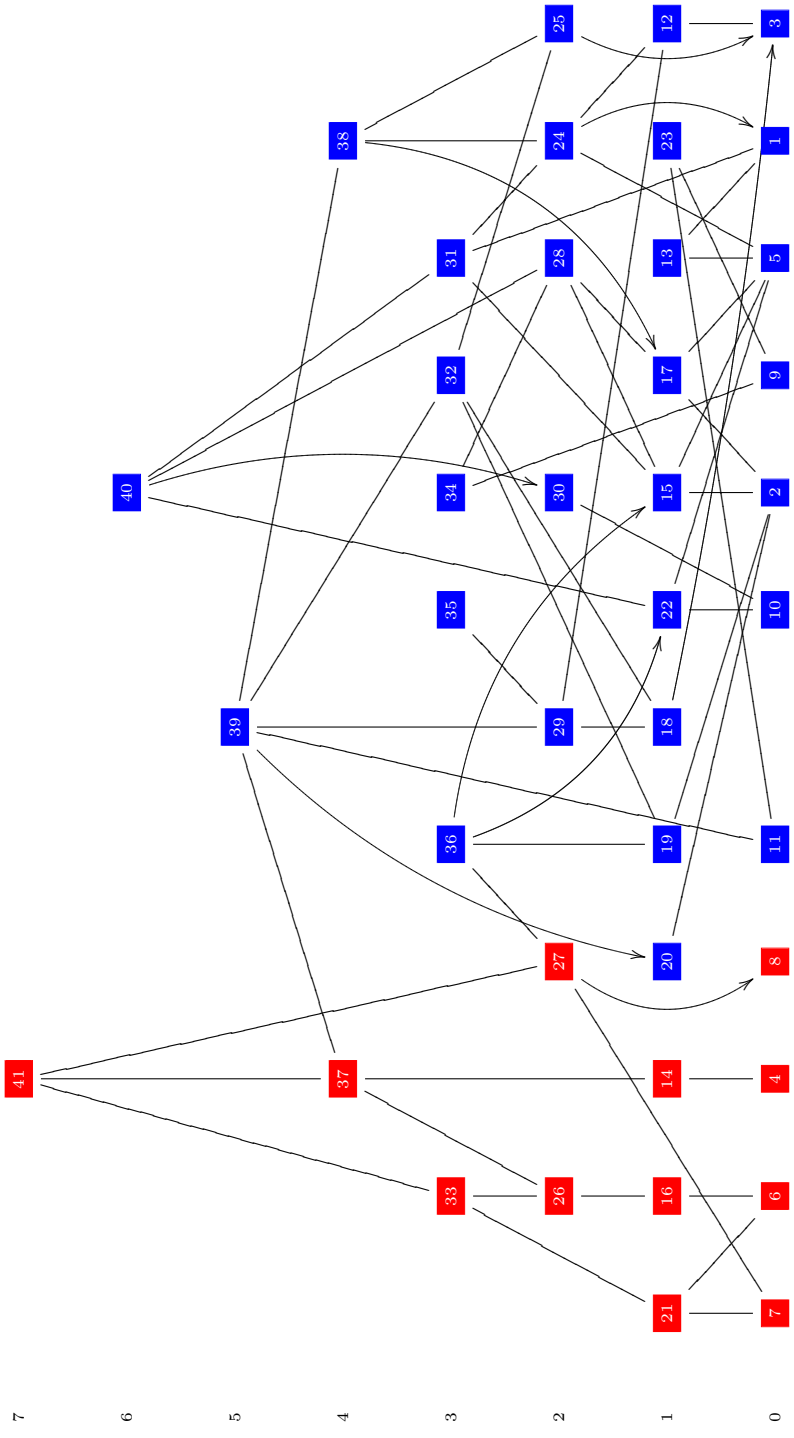
Each of these cases is an irreducible locus in $\mathcal{M}_3$. It is reasonable to know the inclusions between such loci. Such inclusions could help in determining all the cases. We display such inclusions in the following diagram.


$L_3(2)$
S_4
96
D_8
V_4
16
48
S_3
C_2
C_3
C_6
C_9
C_4
16
48
C_2^3
D_{12}
32
C_2
V_4
$C_2 \times C_4$
24
C_{14}
5
4
3
2
1
0

3.1.2. *Genus 4.* In the case of genus $g = 4$ we get the following groups and their corresponding signatures. The Table is provided by K. Magaard.

#	dim	G	ID	sig	type	subs
1	0	S_5	(120,34)	0-$(2,4,5)$	1	
2	0	$C_3 \times S_4$	(72,42)	0-$(2,3,12)$	3	
3	0		(72,40)	0-$(2,4,6)$	4	
4	0	V_{10}	(40,8)	0-$(2,4,10)$	7	
5	0	$C_6 \times S_3$	(36,12)	0-$(2,6,6)$	10	
6	0	U_8	(32,19)	0-$(2,4,16)$	16	
7	0	$SL_2(3)$	(24,3)	0-$(3,4,6)$	20	
8	0	C_{18}	(18,2)	0-$(2,9,18)$	27	
9	0	C_{15}	(15,1)	0-$(3,5,15)$	38	
10	0	C_{12}	(12,2)	0-$(4,6,12)$	45	
11	0	C_{10}	(10,2)	0-$(5,10,10)$	51	
12	1	S_3^2	(36,10)	0-$(2,2,2,3)$	12	3
13	1	S_4	(24,12)	0-$(2,2,2,4)$	18	1, 2
14	1	$C_2 \times D_5$	(20,4)	0-$(2,2,2,5)$	21	4
15	1	$C_3 \times S_3$	(18,3)	0-$(2,2,3,3)$	30	2, 5
16	1	D_8	(16,7)	0-$(2,2,2,8)$	35	6
17	1	$C_2 \times C_6$	(12,5)	0-$(2,2,3,6)$	46	2, 5
18	1	$C_2 \times S_3$	(12,4)	0-$(2,2,3,6)$	41	3
19	1	A_4	(12,3)	0-$(2,3,3,3)$	43	2
20	1	D_{10}	(10,1)	0-$(2,2,5,5)$	49	1
21	1	Q_8	(8,4)	0-$(2,4,4,4)$	59	6, 7
22	1	C_6	(6,2)	0-$(2,6,6,6)$	66	5, 10
23	1	C_5	(5,1)	0-$(5,5,5,5)$	69	9, 11
24	2	D_6	(12,4)	0-(2^5)	40	1, 5, 12
25	2	D_4	(8,3)	0-$(2^4,4)$	57	3, 13
26	2	D_4	(8,3)	0-$(2^4,4)$	56	4, 16
27	2	C_6	(6,2)	0-$(2^3,3,6)$	64	7, 8
28	2	C_6	(6,2)	0-$(2^2,3^3)$	65	15, 17
29	2	S_3	(6,1)	0-$(2^2,3^3)$	62	12, 18
30	2	C_4	(4,1)	0-$(2,4^4)$	77	10
31	3	S_3	(6,1)	0-(2^6)	61	13, 15, 24
32	3	V_4	(4,2)	1-$(2,2,2)$	72	18, 19, 25
33	3	C_4	(4,1)	0-$(2^4,4^2)$	76	21, 26
34	3	C_3	(3,1)	0-(3^6)	80	9, 28
35	3	C_3	(3,1)	0-(3^6)	81	29
36	3	C_3	(3,1)	1-$(3,3,3)$	79	15, 19, 22, 27
37	4	V_4	(4,2)	0-(2^7)	73	14, 26
38	4	V_4	(4,2)	0-(2^7)	74	17, 24, 25
39	5	C_2	(2,1)	2-$(2,2)$	82	11, 20, 29, 32, 37, 38
40	6	C_2	(2,1)	1-(2^6)	83	22, 28, 30, 31, 38
41	7	C_2	(2,1)	0-(2^{10})	84	27, 33, 37

TABLE 5. Hurwitz loci of genus 4 curves


25
12
3
38
24
23
1
31
28
13
5
32
17
9
40
34
30
15
2
35
22
10
39
29
18
36
19
11
27
20
8
41
37
14
4
33
26
16
6
21
7
7
6
5
4
3
2
1
0

In the above diagram are given inclusions among the loci of genus 4 curves. K. Magaard and S. Shpectorov have implemented programs that could compute such inclusions among loci for any reasonable genus.

Problem 5. Determine the inclusions among the loci for all $g \leq 48$

Problem 6. For any group G and a given signature σ such that G occurs as automorphism group of some genus g curve, find the corresponding equation for the curve.

Such problem is open even for small genus. For genus 3 such equations are determined in [18]. However, for $g > 3$ we are unaware of a complete list of equations.

3.2. **Hurwitz curves.** A Hurwitz curve is a genus g curve, defined over an algebraically closed field of characteristic zero, which has $84(g-1)$ automorphisms. A group G that can be realized as an automorphism group of a Hurwitz curve is called a Hurwitz group. There are a lot of papers by group-theoretists on Hurwitz groups, surveyed by Conder. It follows from Hurwitz's presentation that a Hurwitz group is perfect. Thus every quotient is again a Hurwitz group, and if such a quotient is minimal then it is a non-abelian simple group. Several infinite series of simple Hurwitz groups have been found by Conder, Malle, Kuribayashi, Zalessky, Zimmermann and others. In 2001, Wilson showed the monster is a Hurwitz group; see [18] for a complete list of references.

Klein's quartic is the only Hurwitz curve of genus $g \leq 3$. Fricke showed that the next Hurwitz group occurs for $g = 7$ and has order 504. Its group is $SL(2, 8)$, and an equation for it was computed by Macbeath in 1965. Klein's quartic and Macbeath's curve are the only Hurwitz curves whose equations are known. Further Hurwitz curves occur for $g = 14$ and $g = 17$ (and for no other values of $g \leq 19$). It is natural, to try to write equations for these Hurwitz curves of genus 14, 17.

Problem 7. Compute equations for the Hurwitz curves of genus 14, and possibly 17.

4. The automorphism groups of algebraic curves in positive characteristic

The Hurwitz bound is not valid in prime characteristic. Roquette (1970) found that the estimate

$$|G| \leq 84(g-1),$$

on the order of the automorphism group G, holds under the additional assumption $p > g+1$, with one exception: the function field $F = K(x, y)$ with $y^p - y = x^2$ has genus $g = \frac{1}{2}(p-1)$ and $8g(g+1)(2g+1)$ automorphisms.

Stichtenoth (1973) gives a general estimate for the number of automorphisms of a smooth projective curve in characteristic $p > 0$. He proves the inequality

$$|G| < 16 \cdot g^4,$$

but also with one series of exceptions: the function field $F = K(x, y)$ with $y^{p^n} + y = x^{p^{n+1}}$ has genus $g = \frac{1}{2}p^n(p^n - 1)$ and $|G| = p^{3n}(p^{3n} + 1)(p^{2n} - 1)$ automorphisms, so $|G|$ is in this case slightly larger than $16g^4$.

Let X denote a smooth, genus g algebraic curve defined over k, char $k = p > 0$. A theorem of Blichfeld on invariants (in char 0) of subgroups of $PGL_3(k)$ implies

that the genus g curve lifts to characteristic 0 for $p > 2g+1$; see [20, pg. 236-254]. Hence, for large enough p (i.e., $p > 2g+1$) methods described in [18] can be used to determine such groups. Thus, to determine the list of groups that occur as automorphism groups of genus g curves we have to classify the groups that occur for all primes $p \leq 2g+1$.

Since the methods used in [18] are no longer valid in characteristic $p > 0$, a new approach is needed for cases of small characteristic. We suggest the following long term problem:

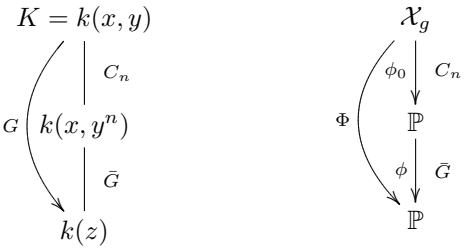
Problem 8. Determine the list of groups which occur as full automorphism groups for genus $g \leq 10$ algebraic curves defined over a field of positive characteristic.

For $g = 2$ this list is well known (it appears also in [34]). However, for $g > 3$ such list of groups is unknown. It would be nice to have a complete list for "small genus", say $g \leq 10$. Since, such lists tend to grow as genus grows, such information could be organized in a database and be very helpful to the mathematics community.

There is a class of curves for which the above problem is a bit easier. These are "cyclic" curves which will be treated next. There is an attempt in [23] to classify all groups which occur as full automorphism groups for genus $g \leq 10$ algebraic curves defined over a field of positive characteristic, including the equations of the curves.

4.1. **Cyclic curves in characteristic $p > 0$.** Let k be an algebraically closed field of characteristic p and $\mathcal{X}_g$ be a genus g cyclic curve defined over k and given by the equation $y^n = f(x)$. Let $K := k(x,y)$ be the function field of $\mathcal{X}_g$. Then $k(x)$ is degree n genus zero subfield of K. Let $G = Aut(K/k)$. Since $C_n := Gal(K/k(x)) = \langle \tau \rangle$, with $\tau^n = 1$ such that $\langle \tau \rangle \lhd G$, then group $\bar{G} := G/C_n$, also $\bar{G} \leq PGL_2(k)$. Hence $\bar{G}$ is isomorphic to one of the following: C_m, D_m, A_4, S_4, A_5, semi direct product of elementary Abelian group with cyclic group, $PSL(2,q)$ and $PGL(2,q)$, see [23].

The group $\bar{G}$ acts on $k(x)$ via the natural way. The fixed field is a genus 0 field, say $k(z)$. Thus z is a degree $|\bar{G}|$ rational function in x, say $z = \phi(x)$. We illustrate with the following diagram:



Let $\phi_0 : \mathcal{X}_g \longrightarrow \mathbb{P}^1$ be the cover which corresponds to the degree n extension $K/k(x)$. Then $\Phi := \phi \circ \phi_0$ has monodromy group $G := Aut(\mathcal{X}_g)$. From the basic covering theory, the group G is embedded in the group S_n where $n = deg\ \Phi$. There is an r-tuple $\overline{\sigma} := (\sigma_1, ..., \sigma_r)$, where $\sigma_i \in S_n$ such that $\sigma_1, ..., \sigma_r$ generate G and $\sigma_1 ... \sigma_r = 1$. The signature of ϕ is an r-tuple of conjugacy classes $\mathbf{C} := (C_1, ..., C_r)$ in S_n such that C_i is the conjugacy class of σ_i. We use the notation n^p to denote the conjugacy class of permutations which are a product of p cycles of length n. Using the signature of $\phi : \mathbb{P}^1 \longrightarrow \mathbb{P}^1$ one finds out the signature of $\Phi : \mathcal{X}_g \longrightarrow \mathbb{P}^1$ for any given g and G.

Let E be the fixed field of G, the Hurwitz genus formula states that

$$2(g_K - 1) = 2(g_E - 1)|G| + deg(\mathfrak{D}_{K/E}) \tag{1}$$

with g_K and g_E the genera of K and E respectively and $\mathfrak{D}_{K/E}$ the different of K/E. Let $\overline{P}_1, \overline{P}_2, ..., \overline{P}_r$ be ramified primes of E. If we set $d_i = deg(\overline{P}_i)$ and let e_i be the ramification index of the $\overline{P}_i$ and let β_i be the exponent of $\overline{P}_i$ in $\mathfrak{D}_{K/E}$. Hence, Eq. (1) may be written as

$$2(g_K - 1) = 2(g_E - 1)|G| + |G| \sum_{i=1}^{r} \frac{\beta_i}{e_i} d_i \tag{2}$$

If $\overline{P}_i$ is tamely ramified then $\beta_i = e_i - 1$ or if $\overline{P}_i$ is wildly ramified then $\beta_i = e_i^* q_i + q_i - 2$ with $e_i = e_i^* q_i$, e_i^* relatively prime to p, q_i a power of p and $e_i^* | q_i - 1$.

Let $\bar{G}$ be a finite subgroup of $PGL_2(q)$ acting on the field $\mathbb{F}_q(x)$. Then, $\bar{G}$ is isomorphic to one of the following groups C_m, D_m, A_4, S_4, A_5, $U = (\mathbb{Z}/p\mathbb{Z})^t$, K_m, $PSL_2(q)$ and $PGL_2(q)$. Then, G is a degree n extension of one of these groups.

Problem 9. Determine the list of groups that occur as full automorphism groups of cyclic curves defined over an algebraically closed field of characteristic $p > 0$.

As stated above, there is an attempt in [23] to completely solve this problem.

5. On the decomposition of Jacobians of algebraic curves with automorphisms

Let $\mathcal{X}$ be a genus g algebraic curve with automorphism group $G := \text{Aut}\ (\mathcal{X})$. Let $H \leq G$ such that $H = H_1 \cup \cdots \cup H_t$ where the subgroups $H_i \leq H$ satisfy $H_i \cap H_j = \{1\}$ for all $i \neq j$. Then,

$$\text{Jac}^{\,t-1}(\mathcal{X}) \times \text{Jac}^{\,|H|}(\mathcal{X}/H) \sim \text{Jac}^{\,|H_1|}(\mathcal{X}/H_1) \times \cdots \text{Jac}^{\,|H_t|}(\mathcal{X}/H_t)$$

The group H satisfying these conditions is called a group with partition. Elementary abelian p-groups, the projective linear groups $PSL_2(q)$, Frobenius groups, dihedral groups are all groups with partition.

Let $H_1, \ldots, H_t \leq G$ be subgroups with $H_i \cdot H_j = H_j \cdot H_i$ for all $i, j \leq t$, and let g_{ij} denote the genus of the quotient curve $\mathcal{X}/(H_i \cdot H_j)$. Then, for $n_1, \ldots, n_t \in \mathbb{Z}$ the conditions

$$\sum n_i n_j g_{ij} = 0, \quad \sum_{j=1}^{t} n_j g_{ij} = 0,$$

imply the isogeny relation

$$\prod_{n_i > 0} \text{Jac}^{\,n_i}(\mathcal{X}/H_i) \sim \prod_{n_j < 0} \text{Jac}^{\,|n_j|}(\mathcal{X}/H_j)$$

In particular, if $g_{ij} = 0$ for $2 \leq i < j \leq t$ and if

$$g = g_{\mathcal{X}/H_2} + \cdots + g_{\mathcal{X}/H_t}$$

then

$$\text{Jac}\ (\mathcal{X}) \sim \text{Jac}\ (\mathcal{X}/H_2) \times \cdots \times \text{Jac}\ (\mathcal{X}/H_t)$$

The reader can check [14, 13] for the proof of the above statements.

Problem 10. Using the structure of automorphism groups of algebraic curves of genus $g \leq 48$, determine possible decompositions of Jacobians for these curves.

Next we focus on the case $g = 3$. This hopefully will give an idea to the reader that such computations are possible.

5.1. **Decomposing Jacobians of genus three algebraic curves with automorphisms.** The inclusions of the loci $\mathcal{M}_3(G, \mathbf{C})$ will help us determine relations in terms of the theta-nulls in each case. We need the following result the proof of each is elementary and we skip the details.

Let $\mathcal{X}$ be a genus 3 curve and $\sigma \in \text{Aut}\,(\mathcal{X})$ an involution. Denote by π the quotient map

$$\pi : \mathcal{X} \to \mathcal{X}/\langle\sigma\rangle$$

The quotient curve $\mathcal{X}/\langle\sigma\rangle$ has genus 0 or 1. If $g(\mathcal{X}/\langle\sigma\rangle) = 0$ then $\mathcal{X}$ is an hyperelliptic curve and σ is the hyperelliptic involution. If $g(\mathcal{X}/\langle\sigma\rangle) = 1$ then σ is called an **elliptic involution**. Then, we have the following.

Lemma 11. *Every involution of a genus 3 non-hyperelliptic curve is an elliptic involution*

Denote by N_i the number of elliptic involutions of a curve $\mathcal{X}$ and the number of conjugacy classes of involutions in Aut $(\mathcal{X})$ by N_c. Both N_i and N_c are displayed for non-hyperelliptic case. We use the information on the automorphism groups to decompose the corresponding Jacobians of curves.

We will use the above facts to decompose the Jacobians of genus 3 non-hyperelliptic curves. $\mathcal{X}$ denotes a genus 3 non-hyperelliptic curve unless otherwise stated and $\mathcal{X}_2$ denotes a genus 2 curve.

5.1.1. *The group C_2.* Then the curve $\mathcal{X}$ has an elliptic involution $\sigma \in \text{Aut}\,(\mathcal{X})$. Hence, there is a Galois covering $\pi\colon \mathcal{X} \to \mathcal{X}/\langle\sigma\rangle =: \mathcal{E}$. We can assume that this covering is maximal. The induced map $\pi^* : \mathcal{E} \to \text{Jac}\,(\mathcal{X})$ is injective. Then, the kernel projection Jac $(\mathcal{X}) \to \mathcal{E}$ is a dimension 2 abelian variety. Hence, there is a genus 2 curve $\mathcal{X}_2$ such that

$$\text{Jac}\,(\mathcal{X}_2) \sim \mathcal{E} \times \text{Jac}\,(\mathcal{X}_2)$$

5.1.2. *The Klein 4-group.* Next, we focus on the automorphism groups G such that $V_4 \hookrightarrow G$. As can be seen from Fig. 1, most groups contain an isomorphic copy of V_4. In this case, there are three elliptic involutions in V_4, namely $\sigma, \tau, \sigma\tau$. Obviously they form a partition. Hence, the Jacobian of $\mathcal{X}$ is the product

$$\text{Jac}\,^2(\mathcal{X}) \sim \mathcal{E}_1^2 \times \mathcal{E}_2^2 \times \mathcal{E}_3^2$$

of three elliptic curves. By applying the Poincare duality we get

$$\text{Jac}\,(\mathcal{X}) \sim \mathcal{E}_1 \times \mathcal{E}_2 \times \mathcal{E}_3$$

5.1.3. *The dihedral group D_8.* In this case, we have 5 involutions in G in 3 conjugacy classes. No conjugacy class has three involutions. Hence, we can pick three involutions such that two of them are conjugate to each other in G and all three of them generate V_4. Hence, Jac $(\mathcal{X}) \sim \mathcal{E}_1^2 \times E_2$, for some elliptic curves $\mathcal{E}_1, \mathcal{E}_2$.

5.1.4. *The symmetric group S_4.* The Jacobian of such curves splits into a product of elliptic curves since $V_4 \hookrightarrow S_4$. Below we give a direct proof of this.

We know that there are 9 involutions in S_4, six of which are transpositions. The other three are product of two 2-cycles and we denote them by $\sigma_1, \sigma_2, \sigma_3$. Let H_1, H_2, H_3 denote the subgroups generated by $\sigma_1, \sigma_2, \sigma_3$. They generate V_4 and are all isomorphic in G. Hence, Jac $(\mathcal{X}) \sim \mathcal{E}^3$, for some elliptic curve $\mathcal{E}$.

5.1.5. *The symmetric group S_3.* We know from above that the Jacobian is a direct product of three elliptic curves. Here we will show that two of those elliptic curves are isomorphic. Let H_1, H_2, H_3 be the subgroups generated by transpositions and H_4 the subgroup of order 3. Then

$$\text{Jac }^3(\mathcal{X}) \sim \mathcal{E}_1^2 \times \mathcal{E}_2^2 \times \mathcal{E}_3^2 \times \text{Jac }^3(\mathcal{Y})$$

for three elliptic curves $\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3$ fixed by involutions and a curve $\mathcal{Y}$ fixed by the element of order 3. Simply by counting the dimensions we have $\mathcal{Y}$ to be another elliptic curve $\mathcal{E}_4$. Since all the transpositions of S_3 are in the same conjugacy class then $\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3$ are isomorphic. Then by applying the Poincare duality we have that

$$\text{Jac }(X) \sim \mathcal{E}^2 \times \mathcal{E}'$$

Summarizing, we have the following:

Theorem 12. *Let $\mathcal{X}$ be a genus 3 curve and G its automorphism group. Then,*
a) If $\mathcal{X}$ is hyperelliptic then

i) If G is isomorphic to V_4 and $C_2 \times C_4$, then Jac (X) is isogenous to the product of and elliptic curve and the Jacobian of a genus 2 curve $\mathcal{X}_2$

$$Jac\ (\mathcal{X}) \sim \mathcal{E} \times Jac\ (\mathcal{X}_2)$$

ii) If G is isomorphic to C_2^3 then Jac (X) is isogenous to the product of three elliptic curves

$$Jac\ (\mathcal{X}) \sim \mathcal{E}_1 \times \mathcal{E}_2 \times \mathcal{E}_3$$

iii) If G is isomorphic to $D_{12}, C_2 \times S_4$ or any of the groups of order 24 or 32, then Jac (X) is isogenous to the product of three elliptic curve such that two of them are isomorphic

$$Jac\ (\mathcal{X}) \sim \mathcal{E}_1^2 \times \mathcal{E}_2$$

b) If $\mathcal{X}$ is non-hyperelliptic then the following hold

i) If G is isomorphic to C_2 then Jac (X) is isogenous to the product of an elliptic curve and the Jacobian of some genus 2 curve $\mathcal{X}_2$

$$Jac\ (\mathcal{X}) \sim \mathcal{E} \times Jac\ (\mathcal{X}_2)$$

ii) If G is isomorphic to V_4 then Jac (X) is isogenous to the product of three elliptic curves

$$Jac\ (\mathcal{X}) \sim \mathcal{E}_1 \times \mathcal{E}_2 \times \mathcal{E}_3$$

iii) If G is isomorphic to S_3, D_8 or has order 16 or 48 then Jac (X) is isogenous to the product of three elliptic curves such that two of them are isomorphic to each other

$$Jac\ (\mathcal{X}) \sim \mathcal{E}_1^2 \times \mathcal{E}_2$$

iv) If G is isomorphic to $S_4, L_3(2)$ or $C_2^3 \rtimes S_3$ then Jac (X) is isogenous to the product of three elliptic curves such that all three of them are isomorphic to each other

$$Jac\ (\mathcal{X}) \sim \mathcal{E}^3.$$

Proof. The proof of the hyperelliptic case is similar and we skip the details. The reader interested in details can check [21].

Part b): When G is isomorphic to C_2, V_4, D_8, S_4, S_3 the result follows from the remarks above. The rest of the theorem is an immediate consequence of Fig. 2. If $|G| = 16, 48$ then $D_8 \hookrightarrow G$. Then, from the remarks at the beginning of this section the results follows. If G is isomorphic to $L_3(2)$ or $C_4^2 \rtimes S_3$ then $S_4 \hookrightarrow G$. Hence the Jacobian splits as in the case of S_4. This completes the proof. □

It is possible that given the equation of $\mathcal{X}$ one can determine the equations of the elliptic or genus 2 components in all cases of the theorem. However, we feel that is outside the scope of this paper.

Problem 13. Determine all algebraic curves of genus $g \leq 10$ such that their Jacobian splits into a product of elliptic curves.

Of course, hyperelliptic curves are an easy exercise to do. There is a little more work for non-hyperelliptic curves since a description of the automorphism group is needed.

6. Invariants of binary forms

It is an interesting and difficult problem in algebraic geometry is to obtain a generalization of the theory of elliptic modular functions to the case of higher genus. In the elliptic case this is done by the so-called *j-invariant* of elliptic curves. In the case of genus $g = 2$, Igusa (1960) gives a complete solution via *absolute invariants* i_1, i_2, i_3 of genus 2 curves. Generalizing such results to higher genus is much more difficult due to the existence of non-hyperelliptic curves. However, even restricted to the hyperelliptic moduli $\mathcal{H}_g$ the problem is still unsolved for $g \geq 3$. In other words, there is no known way of identifying isomorphism classes of hyperelliptic curves of genus $g \geq 3$. In terms of classical invariant theory this means that the field of invariants of binary forms of degree $2g + 2$ is not known for $g \geq 3$.

The following is a special case of $g = 3$.

Problem 14. Find invariants which classify the isomorphism classes of genus 3 hyperelliptic curves.

This is equivalent with determining the field of invariants of binary octavics. The covariants of binary octavics were determined in 1880 by von Gall. The generators of the ring of invariants were determined by Shioda in 1965 where the relations among the $SL_2(\mathbb{C})$ invariants were also determined. However, we believe that such relations are not correct and have been unable to verify them using computational algebra tools.

Problem 15. Determine the ring of invariants of binary octavics.

Other invariants: In a joint paper with J. Gutierrez, we find invariants that identify isomorphism classes of genus g hyperelliptic curves with extra (non-hyperelliptic) involutions; see [11]. This result gives a nice way of doing computations with these curves. We call such invariants *dihedral invariants* of hyperelliptic curves. Let $\mathcal{L}_g$ be the locus in $\mathcal{H}_g$ of hyperelliptic curves with extra involutions. $\mathcal{L}_g$ is a g-dimensional subvariety of $\mathcal{H}_g$. The dihedral invariants yield a birational parametrization of $\mathcal{L}_g$. Computationally these invariants give an efficient way of determining a point of the moduli space $\mathcal{L}_g$.

Dihedral invariants were generalized by Antoniadis and Kontogeorgis to all cyclic curves defined over any algebraically closed field (positive characteristic included); see [4] for details.

Recall that such invariants were defined for "cyclic" curves with extra involutions. Indeed they parameterize the locus of the cyclic curves on the top levels of the diagrams; see genus 3 and 4 cases.

Problem 16. Define similar invariants for all cases of cyclic curves which correspond to higher dimension locus in the moduli $\mathcal{M}_g$. In other words, describe the "cyclic" moduli similar to the hyperelliptic moduli in all cases.

Problem 17. Determine algebraic relations among "dihedral" invariants for all subloci of the "cyclic" moduli.

7. Theta functions of algebraic curves

Let $\pi\colon \mathcal{X}_g \to \mathcal{X}_{g_0}$ be a m-sheeted covering of Riemann surfaces of genus g and g_0, where $g_0 \geq 1$. The general goal is to find properties that $\mathcal{X}_g$ (or rather, the Jacobian of $\mathcal{X}_g$) has, due to the existence of the covering π. This is done by the theta functions of the $\mathcal{X}_g$. This is an old problem that goes back to Riemann and Jacobi. Many other mathematicians have worked on the cases of small genus and small degree, most notably Frobenius, Prym, Königsberger, Rosenhein, Göpel, among others. In [22] we give a historical account of such problems and the significance in modern mathematics.

Let $\mathcal{X}_g$ be an irreducible, smooth, projective curve of genus $g \geq 3$, defined over the complex field $\mathbb{C}$. We denote by $\mathcal{M}_g$ the moduli space of smooth curves of genus g and by Aut $(\mathcal{X}_g)$ the automorphism group of $\mathcal{X}_g$. Each group $G \leq$ Aut $(\mathcal{X}_g)$ acts faithfully on the g-dimensional vector space of holomorphic differential forms on $\mathcal{X}_g$.

The locus of curves in $\mathcal{M}_g$ with fixed automorphism group consists of finitely many components; to determine their number requires mapping class group action on generating systems. We denote by $\mathcal{M}_g(G, \sigma)$ be the sublocus in $\mathcal{M}_g$ of all the genus g curves $\mathcal{X}$ with $G \hookrightarrow$ Aut $(\mathcal{X})$ and signature σ.

Problem 18. Describe the loci $\mathcal{M}_g(G, \sigma)$ in terms of the theta nulls for any given g, G, and σ.

Next we describe in more detail the basic definitions and what is known about this problem.

7.1. **Theta functions and Jacobians of curves.** Let $\mathcal{H}_g$ be the Siegel upper-half space. The symplectic group $Sp(2g, \mathbb{Z})$ acts on $\mathcal{H}_g$ and there is in injection

$$\mathcal{M}_g \hookrightarrow \mathcal{H}_g / SP(2g, \mathbb{Z}) =: \mathcal{A}_g$$

For any $z \in \mathbf{C}^g$ and $\tau \in \mathcal{H}_g$ the **Riemann's theta function** is defined as

$$\theta(z, \tau) = \sum_{u \in \mathbb{Z}^g} e^{\pi i (u^t \tau u + 2u^t z)}.$$

It is holomorphic on $\mathbf{C}^g \times \mathcal{H}_g$ and satisfies

$$\theta(z + u, \tau) = \theta(z, \tau), \quad \theta(z + u\tau, \tau) = e^{-\pi i (u^t \tau u + 2z^t u)} \cdot \theta(z, \tau),$$

where $u \in \mathbb{Z}^g$.

Now let $\mathcal{X}$ be a genus $g \geq 2$ algebraic curve. Choose a symplectic homology basis for $\mathcal{X}$, say

$$\{A_1, \dots A_g, B_1, \dots, B_g\}$$

such that the intersection products $A_i \cdot A_j = B_i \cdot B_j = 0$ and $A_i \cdot B_j = \delta_{ij}$.

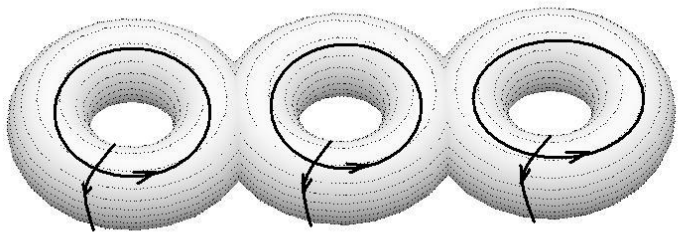


FIGURE 2. A symplectic basis for a genus 3 Riemann surface

We choose a basis $\{w_i\}$ for the space of holomorphic 1-forms such that $\int_{A_i} w_j = \delta_{ij}$. The matrix $\Omega = \left[\int_{B_i} w_j\right]$ is the period matrix of $\mathcal{X}$ and $\Omega \in \mathcal{H}_g$. The columns of the matrix $[I\ |\Omega]$ form a lattice L in $\mathbf{C}^g$ and the Jacobian Jac $(\mathcal{X})$ of $\mathcal{X}$ is Jac $(\mathcal{X}) = \mathbf{C}^g/L$. The **Riemann's theta function** of $\mathcal{X}$ with respect to the above basis is

$$\theta(z, \Omega) = \sum_{u \in \mathbb{Z}^g} e^{\pi i (u^t \Omega u + 2u^t z)},$$

and the locus

$$\Theta := \{z \in \mathbf{C}^g/L : \theta(z, \Omega) = 0\}$$

is called the **theta divisor** of $\mathcal{X}$. Points of order n on Jac $(\mathcal{X})$ are called the $\frac{1}{n}$-**periods**. In the next section we will use the half-periods and quarter-periods to describe the locus of curves in $\mathcal{M}_g$ with fixed automorphism group. For any two half-periods α, β we identify them with their images in $\mathbb{H}_1(\mathcal{X}_g, \mathbb{Z}_2)$, then the **Weil pairing** is defined as

$$|\alpha, \beta| = (-1)^{\alpha \cdot \beta}$$

where $\alpha \cdot \beta$ is the intersection product.

Problem 19. Let G be an automorphism group of a genus $\mathcal{X}_g$ curve and $\mathcal{M}_g(G, \sigma)$ denote the locus of genus g curves with automorphism group G of some signature σ. For $g \geq 4$, describe the locus $\mathcal{M}_g(G, \sigma)$ in terms of the vanishing theta-nulls.

Part 2. Higher dimension varieties

In this part we suggest some problems on higher dimensional varieties. This is by no means a list which includes the most important problems, but simply a list of problems which have special interest to the author.

8. The degree of a rational map

Let k be a field and $\phi : k^n \to k^m$ be a rational map. It is an important problem in algebraic geometry to determine the degree of the map phi. Let us assume that

$$\begin{aligned} \phi : \ & k^n \to k^m \\ & (x_1, \ldots, x_n) \to (f_1, \ldots, f_m) \end{aligned}$$

where $f_1, \ldots, f_m \in k(x_1, \ldots, x_m)$. We assume that $f_i = \frac{p_i(x)}{q_i(x)}$ and $\deg f_i = d_i$, for $i = 1, \ldots, m$. The classical way to determine the degree of such map is as follows: pick a general point $y = (y_1, \ldots, y_m) \in k^m$ such that $\phi(x) = y$. Solve the system of equations

$$\begin{cases} p_1(x) - y_1 q_1(x) = 0 \\ \qquad\qquad\qquad \ldots \\ \qquad\qquad\qquad \ldots \\ p_m(x) - y_m q_m(x) = 0 \end{cases}$$

the number of solutions of such system is bounded by $\Pi_{i=1}^m d_i$. There are some computational issues with this approach though. First, how do we make sure that the point $y \in k^m$ is a generic point. Second, the solution of the above system will involve a Groebener basis argument. Such method is extremely inefficient and will not work well for high degrees.

Problem 20. Combine the symbolic and numerical methods to design an efficient algorithm for determining the degree of a rational map.

There are some attempts to do this by Sommesse et al. However, we are still not aware of how efficient their methods are and if they have been implemented.

9. Parameterizing surfaces

It is a well known fact that if an algebraic curve has genus zero than it can be parameterizable. There are many papers on the parametrization of algebraic curves. The algorithms on parametrization of curves are quite efficient. Furthermore, there are even some results on how to find a "good" parametrization. There are no analogue results for higher dimensional varieties, even though there have been some attempts for algebraic surfaces. The following problem is important theoretically and in applications. [24]

Problem 21. Let $\mathcal{X}$ be a parametric algebraic surface. Design an algorithm which finds a parametrization of $\mathcal{X}$.

While many authors have studied this problem, we are not aware of any results which would do this efficiently.

10. Acknowledgements

Some of the problems and ideas described in this survey come from many discussions with my collaborators, my students, and other colleagues. I would like to thank them all for sharing their knowledge and insight. In particular, I would like to thank K. Magaard for sharing the tables of automorphism groups with me.

References

[1] Robert D. M. Accola, *On the number of automorphisms of a closed Riemann surface*, Trans. Amer. Math. Soc. **131** (1968), 398–408. MR MR0222281 (36 #5333)

[2] ———, *Two theorems on Riemann surfaces with noncyclic automorphism groups.*, Proc. Amer. Math. Soc. **25** (1970), 598–602. MR MR0259105 (41 #3747)

[3] Robert D. M. Accola and Emma Previato, *Covers of tori: genus two*, Lett. Math. Phys. **76** (2006), no. 2-3, 135–161. MR MR2235401 (2007c:14019)

[4] Jannis A. Antoniadis and Aristides Kontogeorgis, *On cyclic covers of the projective line*, Manuscripta Math. **121** (2006), no. 1, 105–130. MR MR2258533 (2007f:14025)

[5] H. Babu and P. Venkataraman, *Group action on genus 3 curves and their Weierstrass points*, Computational aspects of algebraic curves, Lecture Notes Ser. Comput., vol. 13, World Sci. Publ., Hackensack, NJ, 2005, pp. 264–272. MR MR2182045 (2006h:14044)

[6] Walter L. Baily, Jr., *On the automorphism group of a generic curve of genus* > 2, J. Math. Kyoto Univ. **1** (1961/1962), no. 101–108; correction, 325. MR MR0142552 (26 #121)

[7] Thomas Breuer, *Characters and automorphism groups of compact Riemann surfaces*, London Mathematical Society Lecture Note Series, vol. 280, Cambridge University Press, Cambridge, 2000. MR MR1796706 (2002i:14034)

[8] Algebraic curves and their applications, *"http://algcurves.albmath.org/"*, 2007.

[9] Gerhard Frey, *On elliptic curves with isomorphic torsion structures and corresponding curves of genus* 2, Elliptic curves, modular forms, & Fermat's last theorem (Hong Kong, 1993), Ser. Number Theory, I, Int. Press, Cambridge, MA, 1995, pp. 79–98. MR MR1363496 (96k:11067)

[10] GAP, *Groups, algorithms and programming, http://www.gap-system.org/*, 2006.

[11] J. Gutierrez and T. Shaska, *Hyperelliptic curves with extra involutions*, LMS J. Comput. Math. **8** (2005), 102–115 (electronic). MR MR2135032 (2006b:14049)

[12] Jaime Gutierrez, D. Sevilla, and T. Shaska, *Hyperelliptic curves of genus 3 with prescribed automorphism group*, Computational aspects of algebraic curves, Lecture Notes Ser. Comput., vol. 13, World Sci. Publ., Hackensack, NJ, 2005, pp. 109–123. MR MR2182037 (2006j:14038)

[13] E. Kani and M. Rosen, *Idempotent relations and factors of Jacobians*, Math. Ann. **284** (1989), no. 2, 307–327. MR MR1000113 (90h:14057)

[14] Ernst Kani and Michael Rosen, *Idempotent relations among arithmetic invariants attached to number fields and algebraic varieties*, J. Number Theory **46** (1994), no. 2, 230–254. MR MR1269254 (95c:11080)

[15] Robert M. Kuhn, *Curves of genus* 2 *with split Jacobian*, Trans. Amer. Math. Soc. **307** (1988), no. 1, 41–49. MR MR936803 (89f:14027)

[16] C. Maclachlan, *Abelian groups of automorphisms of compact Riemann surfaces*, Proc. London Math. Soc. (3) **15** (1965), 699–712. MR MR0179348 (31 #3596)

[17] ———, *A bound for the number of automorphisms of a compact Riemann surface.*, J. London Math. Soc. **44** (1969), 265–272. MR MR0236378 (38 #4674)

[18] K. Magaard, T. Shaska, S. Shpectorov, and H. Völklein, *The locus of curves with prescribed automorphism group*, Sūrikaisekikenkyūsho Kōkyūroku (2002), no. 1267, 112–141, Communications in arithmetic fundamental groups (Kyoto, 1999/2001). MR MR1954371

[19] K. Magaard, T. Shaska, and H. Völklein, *Genus two curves with degree 5 elliptic subcovers*, Forum Math.

[20] G. A. Miller, H. F. Blichfeldt, and L. E. Dickson, *Theory and applications of finite groups*, Dover Publications Inc., New York, 1961. MR MR0123600 (23 #A925)

[21] J. Paulhus, *Decomposing jacobians of hyperelliptic curves*, preprint.

[22] E. Previato, T. Shaska, and G. S. Wijesiri, *Thetanulls of curves of small genus with automorphisms*, Albanian J. Math. **1** (2007), 253–270.

[23] Sanjeewa R. and Shaska T., *Cyclic curves and their automorphisms*, work in progress.

[24] D. Sevilla and T. Shaska, *Hyperelliptic curves with reduced automorphism group* A_5, Appl. Algebra Engrg. Comm. Comput. **18** (2007), no. 1-2, 3–20. MR MR2280308

[25] T. Shaska, *Curves of genus 2 with* (N, N) *decomposable Jacobians*, J. Symbolic Comput. **31** (2001), no. 5, 603–617. MR MR1828706 (2002m:14023)

[26] ———, *Computational algebra and algebraic curves.*, SIGSAM Bull. **37** (2003), no. 4, 117–124.

[27] ———, *Computational aspects of hyperelliptic curves*, Computer mathematics, Lecture Notes Ser. Comput., vol. 10, World Sci. Publ., River Edge, NJ, 2003, pp. 248–257. MR MR2061839 (2005h:14073)

[28] ———, *Determining the automorphism group of a hyperelliptic curve*, Proceedings of the 2003 International Symposium on Symbolic and Algebraic Computation (New York), ACM, 2003, pp. 248–254 (electronic). MR MR2035219 (2005c:14037)

[29] ———, *Genus 2 fields with degree 3 elliptic subfields*, Forum Math. **16** (2004), no. 2, 263–280. MR MR2039100 (2004m:11097)

[30] ———, *Some special families of hyperelliptic curves*, J. Algebra Appl. **3** (2004), no. 1, 75–89. MR MR2047637 (2005i:14028)

[31] T. Shaska (ed.), *Computational aspects of algebraic curves*, Lecture Notes Series on Computing, vol. 13, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2005, Papers from the conference held at the University of Idaho, Moscow, ID, May 26–28, 2005. MR MR2182657 (2006e:14003)

[32] ———, *Genus two curves covering elliptic curves: a computational approach*, Computational aspects of algebraic curves, Lecture Notes Ser. Comput., vol. 13, World Sci. Publ., Hackensack, NJ, 2005, pp. 206–231. MR MR2182041 (2006g:14051)

[33] ———, *Subvarieties of the hyperelliptic moduli determined by group actions*, Serdica Math. J. **32** (2006), no. 4, 355–374. MR MR2287373 (2007k:14055)

[34] T. Shaska and H. Völklein, *Elliptic subfields and automorphisms of genus 2 function fields*, Algebra, arithmetic and geometry with applications (West Lafayette, IN, 2000), Springer, Berlin, 2004, pp. 703–723. MR MR2037120 (2004m:14047)

[35] Tetsuji Shioda, *On the graded ring of invariants of binary octavics*, Amer. J. Math. **89** (1967), 1022–1046. MR MR0220738 (36 #3790)

Department of Mathematics and Statistics,
Oakland University
Rochester, MI, 48386.
Email: shaska@oakland.edu

Albanian Journal of Mathematics (ISSN: 1930-1235) was founded by T. Shaska in 2007 with the idea to support Albanian mathematicians in Albania and abroad.

The journal is not associated with any government institutions in Albania or any public or private universities in Albania or abroad. The journal does not charge any fees to the authors and has always been an open access journal. The journal supports itself with private donations and voluntary work from its staff. Its main office is in Vlora, Albania.