VOLUME
22/2014
No. 1

ISSN 1804-1388 (Print)
ISSN 2336-1298 (Online)

COMMUNICATIONS IN MATHEMATICS

Available online at
http://cm.osu.cz

UNIVERSITAS OSTRAVIENSIS

Subscriptions
This journal is available for exchange. Information on exchange and/or subscription is available at: Library of the University of Ostrava, Bráfova 7, 701 03 Ostrava, Czech Republic
Website: `http://cm.osu.cz`

The journal is reviewed in Zentralblatt für Mathematik and Mathematical Reviews.

ISSN 1804-1388 (Print), ISSN 2336-1298 (Online)

Communications in Mathematics 22 (2014) 1–12

Fixed point theorems of G-fuzzy contractions in fuzzy metric spaces endowed with a graph

Satish Shukla

Abstract. Let $(X, M, *)$ be a fuzzy metric space endowed with a graph G such that the set $V(G)$ of vertices of G coincides with X. Then we define a G-fuzzy contraction on X and prove some results concerning the existence and uniqueness of fixed point for such mappings. As a consequence of the main results we derive some extensions of known results from metric into fuzzy metric spaces. Some examples are given which illustrate the results.

1 Introduction

The concept of fuzzy sets was introduced by Zadeh [12]. He considered the nature of uncertainty in the behaviour of systems possessing fuzzy nature by means of a fuzzy set. The concept of fuzzy metric space was introduced by Kramosil and Michálek [7]. George and Veeramani [1] modified the definition of fuzzy metric spaces due to Kramosil and Michálek. The fixed point theory in fuzzy metric spaces was started by Grabiec [13] which has become of interest for several authors. Gregori and Sapena [15] introduced the concept of fuzzy contractive mappings and proved some fixed point results for fuzzy contractive mappings.

On the other hand, Jachymski [11] introduced the fixed point theory in the spaces endowed with a graph. The fixed point results on the spaces endowed with a graph generalize and unify several known results in the literature, e.g., the fixed point results on the spaces endowed with a partial order [3], [8], [10] and the fixed point results for the cyclic mappings (see [6] and [11]).

In this paper, we introduce the G-fuzzy contractions as an extension of Banach G-contraction (see [11]) in fuzzy metric spaces and prove some fixed point results for such mappings in complete fuzzy metric spaces in the sense of Grabiec [13]. Our results are the extension of results of Jachymski [11] and a generalization of result of Gregori and Sapena [15] in fuzzy metric spaces.

2010 MSC: 54H25, 47H10
Key words: graph, partial order, fuzzy metric space, contraction, fixed point

2 Preliminaries

Firstly, we recall some known definitions and the properties about the fuzzy metric spaces.

Definition 1 (Schweizer and Sklar [4]). A binary operation $T\colon [0,1]\times[0,1]\to[0,1]$ is called a t-norm if the following conditions are satisfied:

(T1) $T(a,b)=T(b,a)$;

(T2) $T(a,b)\le T(c,d)$ for $a\le c, b\le d$;

(T3) $T(T(a,b),c)=T(a,T(b,c))$;

(T4) $T(a,0)=0, T(a,1)=1$;

for all $a,b,c,d\in[0,1]$.

For $a,b\in[0,1]$, instead of $T(a,b)$ we will use the infix notation $a*b$. For $a_1,a_2,\dots,a_n\in[0,1]$ and $n\in\mathbb{N}$, the product $a_1*a_2*\cdots*a_n$ will be denoted by $\prod_{i=1}^n a_i$. For the details concerning t-norms the reader is referred to [5], [14].

In the present paper we will use the following definition of a fuzzy metric space:

Definition 2 (George and Veeramani [1]). A triple $(X,M,*)$ is called a fuzzy metric space if X is a nonempty set, $*$ is a continuous t-norm and $M\colon X^2\times(0,\infty)\to[0,1]$ is a fuzzy set satisfying following conditions:

(GV1) $M(x,y,t)>0$;

(GV2) $M(x,y,t)=1$ if and only if $x=y$;

(GV3) $M(x,y,t)=M(y,x,t)$;

(GV4) $M(x,z,t+s)\ge M(x,y,t)*M(y,z,s)$;

(GV5) $M(x,y,\cdot)\colon(0,\infty)\to[0,1]$ is a continuous mapping;

for all $x,y,z\in X$ and $s,t>0$.

Example 1 (George and Veeramani [1]). Let (X,d) be a metric space, then the triple $(X,M_d,*)$ is a fuzzy metric space, where $a*b=ab$ for all $a,b\in[0,1]$ and

$$M_d(x,y,t)=\frac{t}{t+d(x,y)}\text{ for all } x,y\in X, t>0.$$

M_d is called the standard fuzzy metric induced by the metric d.

Let $(X,M,*)$ be a fuzzy metric space. An open ball $B(x,r,t)$ with center $x\in X$ and radius r, $0<r<1$ and $t>0$ is defined by

$$B(x,r,t)=\{y\in X\colon M(x,y,t)>1-r\}.$$

The collection $\{B(x,r,t)\colon x\in X, 0<r<1, t>0\}$ is a neighbourhood system for the topology τ on X induced by the fuzzy metric M.

For topological properties of a fuzzy metric space in the sense of George and Veeramani the reader is referred to [1].

Remark 1 (George and Veeramani [2]). Let $(X, M, *)$ be a fuzzy metric space, then the function $M(x, y, \cdot)$ is a nondecreasing function.

Theorem 1 (George and Veeramani [1]). *Let $(X, M, *)$ be a fuzzy metric space, and τ be the topology induced by the fuzzy metric. Then for a sequence $\{x_n\}$ in X, $x_n \to x$ if and only if*

$$\forall_{t>0} \quad \lim_{n\to\infty} M(x_n, x, t) = 1\,.$$

In this paper, we use the following definitions of Cauchy sequence and complete fuzzy metric space.

Definition 3 (Grabiec [13]). Let $(X, M, *)$ be a fuzzy metric space and $\{x_n\}$ be a sequence in X. Then $\{x_n\}$ is called a Cauchy sequence if

$$\forall_{t>0} \ \forall_{p>0} \quad \lim_{n\to\infty} M(x_{n+p}, x_n, t) = 1\,.$$

A complete fuzzy metric space is a fuzzy metric space in which every Cauchy sequence is convergent.

Definition 4 (Gregori and Sapena [15]). Let $(X, M, *)$ be a fuzzy metric space. A mapping $T\colon X \to X$ is called t-uniformly continuous if for all $r \in (0, 1)$ there exists $s \in (0, 1)$ such that

$$\forall_{x,y\in X} \ \forall_{t>0} \quad \big[M(x, y, t) \geq 1 - s \ \ \Rightarrow \ \ M(Tx, Ty, t) \geq 1 - r\big]\,.$$

Remark 2. If T is t-uniformly continuous then it is uniformly continuous for the uniformity generated by M, thus it is continuous for the topology deduced from M. For the details concerning a uniform structure in a fuzzy metric space, see [15].

Definition 5 (Gregori and Sapena [15]). Let $(X, M, *)$ be a fuzzy metric space. A mapping $T\colon X \to X$ is called a fuzzy contractive mapping if there exists $\lambda \in (0, 1)$ such that

$$\forall_{x,y\in X} \ \forall_{t>0} \quad \frac{1}{M(Tx, Ty, t)} - 1 \leq \lambda \left[\frac{1}{M(x, y, t)} - 1\right]. \qquad (1)$$

It is obvious that if T is a fuzzy contractive mapping then it is t-uniformly continuous and so continuous.

Following concepts about the graphs are similar to those in [11].

Let $(X, M, *)$ be a fuzzy metric space. Let Δ denote the diagonal of the Cartesian product $X \times X$. Consider a directed graph G such that the set $V(G)$ of its vertices coincides with X, and the set $E(G)$ of its edges contains all loops, i.e., $E(G) \supseteq \Delta$. We assume G has no parallel edges, so we can identify G with the pair $(V(G), E(G))$. Moreover, we may treat G as a weighted graph by assigning to each edge the fuzzy distance between its vertices.

By G^{-1} we denote the conversion of a graph G, i.e., the graph obtained from G by reversing the direction of edges. Thus we have

$$E(G^{-1}) = \{(x, y) \in X \times X\colon (y, x) \in E(G)\}\,.$$

The letter $\widetilde{G}$ denotes the undirected graph obtained from G by ignoring the direction of edges. Actually, it will be more convenient for us to treat $\widetilde{G}$ as a directed graph for which the set of its edges is symmetric. Under this convention,

$$E(\widetilde{G}) = E(G) \cup E(G^{-1}). \tag{2}$$

If x and y are vertices in a graph G, then a path in G from x to y of length l is a sequence $(x_i)_{i=0}^{l}$ of $l+1$ vertices such that $x_0 = x, x_l = y$ and $(x_{i-1}, x_i) \in E(G)$ for $i = 1, \dots, l$. A graph G is called connected if there is a path between any two vertices of G. A graph G is weakly connected if $\widetilde{G}$ is connected. For a graph G such that $E(G)$ is symmetric and x is a vertex in G, the subgraph G_x consisting of all edges and vertices which are contained in some path beginning at x is called the component of G containing x. In this case $V(G_x) = [x]_G$, where $[x]_G$ is the equivalence class of a relation R defined on $V(G)$ by the rule: yRz if there is a path in G from y to z. Clearly, G_x is connected.

Now we can state our main results.

3 Main results

Throughout this section we assume that X is nonempty set, G is a directed graph such that $V(G) = X$ and $E(G) \supseteq \Delta$.

First we define the Cauchy equivalent sequence and G-fuzzy contraction in fuzzy metric spaces.

Definition 6. Let $(X, M, *)$ be a fuzzy metric space and G be a graph. Two sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ in X are said to be Cauchy equivalent if each of them is a Cauchy sequence and $\lim_{n\to\infty} M(x_n, y_n, t) = 1$ for all $t > 0$.

Definition 7. Let $(X, M, *)$ be a fuzzy metric space and G be a graph. The mapping $T\colon X \to X$ is said to be a G-fuzzy contraction if the following conditions hold:

(GF1) $\forall_{x,y\in X}\ \big((x,y) \in E(G) \Rightarrow (Tx, Ty) \in E(G)\big)$, i.e., T is edge-preserving;

(GF2) $$\exists_{\lambda\in(0,1)}\ \forall_{x,y\in X}\ \forall_{t>0}\ \left((x,y)\in E(G) \Rightarrow \frac{1}{M(Tx,Ty,t)} - 1 \le \lambda\left[\frac{1}{M(x,y,t)} - 1\right]\right),$$

where λ is called the contractive constant of T.

An obvious consequence of symmetry of $M(\cdot,\cdot,t)$ and (2) is the following remark.

Remark 3. If T is a G-fuzzy contraction then it is both a G^{-1}-fuzzy contraction and a $\widetilde{G}$-fuzzy contraction.

Example 2. Any constant function $T\colon X \to X$, that is $Tx = c$, $x \in X$, where $c \in X$ is fixed, is a G-fuzzy contraction with arbitrary value of $\lambda \in (0,1)$ since $E(G)$ contains all the loops.

Example 3. Any fuzzy contractive mapping is a G_0-fuzzy contraction with the same contractive constant, where the graph G_0 is defined by $E(G_0) = X \times X$.

Example 4. Let (X, d) be a metric space endowed with a partial order $\sqsubseteq$ and $T\colon X \to X$ be an ordered contraction, i.e.,

$$\exists_{\lambda \in (0,1)} \ \forall_{x,y \in X} \ \left(x \sqsubseteq y \ \Rightarrow \ d(Tx, Ty) \leq \lambda d(x, y)\right).$$

Then T is a G_d-fuzzy contraction in the induced fuzzy metric space $(X, M_d, *)$ with contractive constant λ, where $G_d = \{(x, y) \in X \times X : x \sqsubseteq y\}$.

We see that every fuzzy contractive mapping is t-uniformly continuous. Following example shows that a G-fuzzy contraction need not be even continuous.

Example 5. Let $(\mathbb{R}^+, d)$ be the usual metric space of positive reals and $(\mathbb{R}^+, M_d, *)$ be the standard fuzzy metric space induced by d. Let G be the graph defined by $V(G) = X$ and

$$E(G) = \Delta \cup \{(x, y) \in X \times X : x, y \in \mathbb{Q} \cap \mathbb{R}^+ \text{ with } x \leq y\}$$

Let the mapping $T\colon X \to X$ be defined by

$$Tx = \begin{cases} \dfrac{x}{2}, & \text{if } x \in \mathbb{Q} \cap \mathbb{R}^+; \\ 0, & \text{otherwise.} \end{cases}$$

Then it is clear that T is not continuous. Now one can see easily that T is a G-fuzzy contraction with $\lambda = \frac{1}{2}$.

Definition 8. Let $(X, M, *)$ be a fuzzy metric space and $T\colon X \to X$ be a mapping. We denote the nth iterate of T on $x \in X$ by $T^n x$ and $T^n x = TT^{n-1}x$ for all $n \in \mathbb{N}$ with $T^0 x = x$. T is called a Picard operator if T has a unique fixed point u and $\lim_{n\to\infty} M(T^n x, u, t) = 1$ for all $x \in X, t > 0$. T is called a weakly Picard operator if for all $x \in X$ there exists a fixed point $u_x \in X$ (which may depend on x) of T such that $\lim_{n\to\infty} M(T^n x, u_x, t) = 1$ for all $t > 0$.

Note that every Picard operator is a weakly Picard operator. Also, the fixed point of a weakly Picard operator may not be unique. In further discussion, we will denote the set of all fixed points of T by $\operatorname{Fix} T$. A subset $A \subset X$ is said to be T-invariant if $T(A) \subset A$.

The following lemma will be useful in sequel.

Lemma 1. *Let $T\colon X \to X$ be a G-fuzzy contraction, then given $x \in X$ and $y \in [x]_{\widetilde{G}}$, we have $\lim_{n\to\infty} M(T^n x, T^n y, t) = 1$ for all $t > 0$.*

Proof. Let $x \in X$ and $y \in [x]_{\widetilde{G}}$. Then by definition there exists a path $(x_i)_{i=0}^m$ in $\widetilde{G}$ from x to y, i.e., $x_0 = x$, $x_m = y$ and $(x_i, x_{i-1}) \in E(\widetilde{G})$ for $i = 1, 2, \ldots, m$. By Remark 3, T is a $\widetilde{G}$-fuzzy contraction. Therefore by (GF1) we have $(T^n x_i, T^n x_{i-1}) \in E(\widetilde{G})$ and by (GF2), for $i = 1, 2, \ldots, m$ and $t > 0$ we have

$$\frac{1}{M(T^n x_{i-1}, T^n x_i, t)} - 1 \leq \lambda^n \left[\frac{1}{M(x_{i-1}, x_i, t)} - 1\right]. \tag{3}$$

Now we can choose a strictly decreasing sequence $(a_n)_{n\in\mathbb{N}}$ of positive numbers such that $\sum_{i=1}^{\infty} a_i = 1$ and then using (3) we obtain

$$\begin{aligned} M(T^n x, T^n y, t) &= M\Big(T^n x_0, T^n x_m, \sum_{i=1}^{\infty} a_i t\Big) \\ &\geq M\Big(T^n x_0, T^n x_m, \sum_{i=1}^{m} a_i t\Big) \geq \prod_{i=1}^{m} M(T^n x_{i-1}, T^n x_i, a_i t) \\ &\geq \prod_{i=1}^{m} \left[\frac{1}{1 - \lambda^n + \frac{\lambda^n}{M(x_{i-1}, x_i, a_i t)}}\right]. \end{aligned}$$

As $\lambda \in (0, 1)$ we obtain $\lim_{n\to\infty} M(T^n x, T^n y, t) = 1$ for all $t > 0$. □

The following theorem shows the equivalency of connectedness of graph and the convergence of an iterative sequences in fuzzy metric spaces.

Theorem 2. *The following statements are equivalent:*

(i) *G is weakly connected;*

(ii) *for any G-fuzzy contraction $T\colon X \to X$, given $x, y \in X$ the sequences $(T^n x)_{n\in\mathbb{N}}$ and $(T^n y)_{n\in\mathbb{N}}$ are Cauchy equivalent;*

(iii) *for any G-fuzzy contraction $T\colon X \to X$, $\operatorname{card}(\operatorname{Fix} T) \leq 1$.*

Proof. (i)⇒(ii): Let T be a G-fuzzy contraction and $x, y \in X$ then by hypothesis G is weakly connected, therefore $[x]_{\widetilde{G}} = X$ and so $T^p x \in [x]_{\widetilde{G}}$ for all $p \in \mathbb{N}$. Now by Lemma 1, we have $(T^n x)_{n\in\mathbb{N}}$ is a Cauchy sequence. Similarly, $(T^n y)_{n\in\mathbb{N}}$ is a Cauchy sequence. Since $[x]_{\widetilde{G}} = X$ therefore by Lemma 1, we have $\lim_{n\to\infty} M(T^n x, T^n y, t) = 1$ for all $t > 0$. Hence the sequences $(T^n x)_{n\in\mathbb{N}}$ and $(T^n y)_{n\in\mathbb{N}}$ are Cauchy equivalent.

(ii)⇒(iii): Let $x, y \in \operatorname{Fix} T$, where T is a G-fuzzy contraction. Since $x, y \in \operatorname{Fix} T^n$ and we have $M(x, y, t) = M(T^n x, T^n y, t)$. So by assumption $x = y$.

(iii)⇒(i): Suppose (iii) holds but G is not weakly connected, i.e., $\widetilde{G}$ is disconnected. Let $u \in X$, then both the sets $[u]_{\widetilde{G}}$ and $X \setminus [u]_{\widetilde{G}}$ are nonempty. Let $v \in X \setminus [u]_{\widetilde{G}}$ and define a mapping $T\colon X \to X$ by

$$Tx = \begin{cases} u, & \text{if } \; x \in [u]_{\widetilde{G}}; \\ v, & \text{if } \; x \in X \setminus [u]_{\widetilde{G}}. \end{cases}$$

Now clearly $\operatorname{Fix} T = \{u, v\}$. We show that T is a G-fuzzy contraction. If $(x, y) \in E(G)$ then by the definition we have $[x]_{\widetilde{G}} = [y]_{\widetilde{G}}$, so either $x, y \in [u]_{\widetilde{G}}$ or $u, v \in X \setminus [u]_{\widetilde{G}}$. In both the cases we have $Tx = Ty$ and so $(Tx, Ty) \in E(G)$ (since $E(G) \supseteq \Delta$) and (GF1) is satisfied. Also, $M(Tx, Ty, t) = 1$ for all $t > 0$ so (GF2) is satisfied. Thus T is a G-fuzzy contraction and $\operatorname{card}(\operatorname{Fix} T) = 2 > 1$. This contradiction proves the result. □

The following corollary is an immediate consequence of the above theorem.

Corollary 1. *Let $(X, M, *)$ be a complete fuzzy metric space. Then the following statements are equivalent:*

(i) G is weakly connected;

(ii) for any G-fuzzy contraction $T\colon X \to X$, there is $x^* \in X$ such that $\lim_{n\to\infty} T^n x = x^*$ for all $x \in X$.

The proof of following proposition is similar as for the metric case (see, e.g., [11]).

Proposition 1. Assume that $T\colon X \to X$ is a G-fuzzy contraction such that for some $x_0 \in X$ we have $Tx_0 \in [x_0]_{\widetilde{G}}$. Let $\widetilde{G}_{x_0}$ be the component of $\widetilde{G}$ containing x_0. Then $[x_0]_{\widetilde{G}}$ is T-invariant and $T|_{[x_0]_{\widetilde{G}}}$ is a $\widetilde{G}_{x_0}$-fuzzy contraction. Moreover, if $x, y \in [x_0]_{\widetilde{G}}$, then the sequences $(T^n x)_{n\in\mathbb{N}}$ and $(T^n y)_{n\in\mathbb{N}}$ are Cauchy equivalent.

Definition 9. Let $(X, M, *)$ be a fuzzy metric space and G be a directed graph, $T\colon X \to X$ be a mapping and $x, x^* \in X$. Then we say that the 4-tuple $(X, M, *, G)$ have the property $(\mathcal{P}_T)$ if for any sequence $(T^n x)_{n\in\mathbb{N}}$, which converges to x^* with $(T^n x, T^{n+1} x) \in E(G)$ for all $n \in \mathbb{N}$ there exists is a subsequence $(T^{k_n} x)_{n\in\mathbb{N}}$ with $(T^{k_n} x, x^*) \in E(G)$ for $n \in \mathbb{N}$.

Theorem 3. *Let $(X, M, *)$ be a complete fuzzy metric space and G be a directed graph and let the 4-tuple $(X, M, *, G)$ have the property $(\mathcal{P}_T)$. Let $T\colon X \to X$ be a G-fuzzy contraction and $X_T = \{x \in X\colon (x, Tx) \in E(G)\}$, then the following statements hold:*

(A) if $x \in X_T$, then $T|_{[x]_{\widetilde{G}}}$ is a Picard operator;

(B) if $X_T \neq \emptyset$ and G is weakly connected, then T is a Picard operator;

(C) $\operatorname{Fix} T \neq \emptyset$ if and only if $X_T \neq \emptyset$;

(D) if $T \subseteq E(G)$, then T is a weakly Picard operator.

Proof. To prove (A) let $x \in X_T$. By definition of X_T, $(x, Tx) \in E(G)$ and so we have $Tx \in [x]_{\widetilde{G}}$. Now by Proposition 1, we have $T\colon [x]_{\widetilde{G}} \to [x]_{\widetilde{G}}$ and T is a $\widetilde{G}_x$-fuzzy contraction and if $y \in \widetilde{G}_x$ then $(T^n x)_{n\in\mathbb{N}}$ and $(T^n y)_{n\in\mathbb{N}}$ are Cauchy equivalent and so $(T^n x)_{n\in\mathbb{N}}$ is a Cauchy sequence. By completeness of X and Theorem 1 there exists $x^* \in X$ such that

$$\lim_{n\to\infty} M(T^n x, x^*, t) = 1 \quad \text{for all} \quad t > 0\,. \tag{4}$$

Since $(x, Tx) \in E(G)$ we have $(x, Tx) \in E(\widetilde{G})$ and so by (GF1) we have

$$(T^n x, T^{n+1} x) \in E(G) \quad \text{for all} \quad n \in \mathbb{N}. \tag{5}$$

Now by property $(\mathcal{P}_T)$ there exists a subsequence $(T^{k_n} x)_{n\in\mathbb{N}}$ such that $(T^{k_n} x, x^*) \in E(G)$ for all $n \in \mathbb{N}$. Hence, $(x, Tx, T^2 x, \dots, T^{k_n} x, x^*)$ is a path in G and so in $\widetilde{G}$. Therefore, $x^* \in [x]_{\widetilde{G}}$. Using (GF2) we have

$$\frac{1}{M(T^{k_n+1}x, Tx^*, t)} - 1 \leq \lambda \left[\frac{1}{M(T^{k_n} x, x^*, t)} - 1\right]$$

for all $t > 0$. Using the above inequality we obtain

$$\begin{aligned} M(x^*, Tx^*, t) &\geq M(x^*, T^{k_n+1}x, t/2) * M(T^{k_n+1}x, Tx^*, t/2) \\ &\geq M(x^*, T^{k_n+1}x, t/2) * \left[\frac{1}{1 - \lambda + \frac{\lambda}{M(T^{k_n}x, x^*, t/2)}}\right]. \end{aligned}$$

Letting $n \to \infty$ and using (4) in the above inequality we obtain $M(x^*, Tx^*, t) = 1$ for all $t > 0$. Thus $Tx^* = x^*$, i.e., $x^* \in [x]_{\widetilde{G}}$ is a fixed point of T and so by Theorem 2, $T|_{[x]_{\widetilde{G}}}$ is a Picard operator.

To prove (B) let $X_T \neq \emptyset$ and G is weakly connected then $[x]_{\widetilde{G}} = X$ for all $x \in X_T$ and so by (A) T is a Picard operator.

To prove (C), note that if $\operatorname{Fix} T \neq \emptyset$ then there is some $x \in \operatorname{Fix} T$ then $Tx = x$ and $E(G) \supseteq \Delta$ we have $(x, Tx) \in E(G)$. So $x \in X_T$ and $\operatorname{Fix} T \subseteq X_T \neq \emptyset$. If $X_T \neq \emptyset$, then by (A) for any $x \in X_T$, $T|_{[x]_{\widetilde{G}}}$ is a Picard operator and so $\operatorname{Fix} T \neq \emptyset$.

To prove (D) if $T \subseteq E(G)$, then $(x, Tx) \in E(G)$ for all $x \in X$ and so $X = X_T$. Now the result follows from (A). □

In the above theorem, if $x \in X_T$ then $T|_{[x]_{\widetilde{G}}}$ is a Picard operator, but if G is not weakly connected then T need not be a Picard operator on X, i.e., the fixed point of T need not be unique. The following example illustrates the above Theorem.

Example 6. Let $X = \left\{\frac{1}{2^n} : n \in \mathbb{N}\right\} = X_o \cup X_e$, where $X_o = \left\{\frac{1}{2^n} : n \in \mathbb{N}_o\right\}$, $X_e = \left\{\frac{1}{2^n} : n \in \mathbb{N}_e\right\}$ and $\mathbb{N}_o$, $\mathbb{N}_e$ are the set of all odd and even natural numbers respectively. Let $*$ be the product norm, i.e., $a * b = ab$ for all $a, b \in [0, 1]$. Define the fuzzy set $M\colon X^2 \times (0, \infty) \to [0, 1]$ by

$$M(x, y, t) = \begin{cases} 1, & \text{if } x = y; \\ xy, & \text{otherwise} \end{cases} \quad \forall t > 0.$$

Let $T\colon X \to X$ be a mapping defined by

$$T\left(\frac{1}{2^n}\right) = \begin{cases} \frac{1}{2}, & \text{if } x \in \mathbb{N}_o; \\ \frac{1}{4}, & \text{if } x \in \mathbb{N}_e. \end{cases}$$

Let G be the graph with $V(G) = X$ and

$$E(G) = (X_o \times X_o) \cup (X_e \times X_e).$$

Then it is easy to see that T is a G-fuzzy contraction with arbitrary $\lambda \in (0,1)$ and by definition of T the condition $(\mathcal{P}_T)$ holds. Note that for all $k \in \mathbb{N}_o$ we have $\frac{1}{2^k} \in X_T$ and $\left[\frac{1}{2^k}\right]_{\widetilde{G}} = X_o$ and $T|_{X_o}$ is a Picard operator. Similarly, $\frac{1}{2^k} \in X_T$ and $\left[\frac{1}{2^k}\right]_{\widetilde{G}} = X_e$ for all $k \in \mathbb{N}_e$ and $T|_{X_e}$ is a Picard operator.

Now it is easy to see that G is not weakly connected and T is not a Picard operator on X since $\operatorname{Fix} T = \left\{\frac{1}{2}, \frac{1}{4}\right\}$. Also, $T \subseteq E(G)$ and T is a weakly Picard operator on X.

The next example shows that the results of this paper generalize the corresponding classical concepts in the classical metric space.

Example 7. Let $X = \left\{\frac{1}{2^{2^n}} : n \in \mathbb{N}_0\right\}$, $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. Then the triple $(X, M_d, *)$ is a fuzzy metric space, where $a * b = ab$ for all $a, b \in [0,1]$ and

$$M(x,y,t) = \begin{cases} 1, & \text{if } x = y\,; \\ xy, & \text{otherwise} \end{cases} \qquad \text{for all } t > 0\,.$$

Note that there exists no metric d on X satisfying $M(x,y,t) = \dfrac{t}{t + d(x,y)}$. Therefore, this fuzzy metric is not a standard fuzzy metric induced by a metric (in the sense of George and Veeramani [1]). Define a mapping $T \colon X \to X$ by

$$T\left(\frac{1}{2^{2^n}}\right) = \begin{cases} \dfrac{1}{2^{2^{n-1}}}, & \text{if } n \in \mathbb{N}\,; \\ \dfrac{1}{2}, & \text{if } n = 0\,. \end{cases}$$

Let G be the graph with $V(G) = X$ and

$$E(G) = \{(x,y) \in X \times X : x \le y\}\,.$$

Then it is easy to see that T is a G-fuzzy contraction with $\lambda \in [1/2, 1)$. Also, the property $(\mathcal{P}_T)$ is satisfied trivially and $X_T \neq \emptyset$. By definition, the graph G is weakly connected and by (B) of Theorem 3, T is a Picard operator with $\operatorname{Fix} T = \left\{\frac{1}{2}\right\}$.

On the other hand, T is not a Banach contraction with respect to the usual metric d, and therefore it is not a fuzzy contractive mapping with respect to the standard fuzzy metric $M(x,y,t) = \frac{t}{t+d(x,y)}$ induced by d. To see this, take the points $x = \frac{1}{4}, y = \frac{1}{16} \in X$ and then T fails to be a Banach contraction with respect to d.

Now we give some consequences of Theorem 3. The following corollary is the fuzzy metric version and an improvement of the result of Nieto and Rodríguez-López [9].

Corollary 2. *Let $(X, M, *)$ be a complete fuzzy metric space and $\preceq$ be a partial order defined on X. Let $T\colon X \to X$ be a nondecreasing mapping (i.e., $x \preceq y \Rightarrow Tx \preceq Ty$) such that the following contractive condition is satisfied:*

$$\exists_{\lambda\in(0,1)}\ \forall_{x,y\in X}\ \forall_{t>0}\ \left(x \preceq y\ \Rightarrow\ \frac{1}{M(Tx,Ty,t)} - 1 \le \lambda\left[\frac{1}{M(x,y,t)} - 1\right]\right).$$

Assume that the following condition holds:

if there is a nondecreasing sequence $(x_n)_{n\in\mathbb{N}}$ in X which converges to $x \in X$ and $x_{n+1} \preceq x_n$ for all $n \in \mathbb{N}$, then $x_n \preceq x$ or $x \preceq x_n$ for all $n \in \mathbb{N}$. $(\mathcal{P}')$

If there exists $x_0 \in X$ such that $x_0 \preceq Tx_0$ or $Tx_0 \preceq x_0$, then T has a fixed point in X.

Proof. Let G be the graph defined by $V(G) = X$ and

$$E(G) = \{(x, y) \in X \times X : x \preceq y \vee y \preceq x\}.$$

Then since T is nondecreasing (GF1) holds and by the contractive condition (GF2) also holds. Therefore T is a G-fuzzy contraction. Also $(\mathcal{P}')$ implies $(\mathcal{P}_T)$ and by assumption $(x_0, Tx_0) \in E(G)$ so $x_0 \in X_T$. Therefore by (A) of Theorem 3, $T|_{[x_0]_{\tilde{G}}}$ is a Picard operator and so has a fixed point in $T|_{[x_0]_{\tilde{G}}}$. □

Recently, Kirk et al. [16] introduced the idea of cyclic contractions and established fixed point results for such mappings.

Let X be a nonempty set, m a positive integer, $A_i, i = 1, 2, \dots, m$ are nonempty subsets of X and $T\colon \bigcup_{i=1}^m A_i \to \bigcup_{i=1}^m A_i$ be a mapping, then $B = \bigcup_{i=1}^m A_i$ is said to be a cyclic representation of B with respect to T if

$$T(A_1) \subset A_2\,,\ T(A_2) \subset A_3\,,\ \dots\,,\ T(A_m) \subset T(A_1)$$

and then T is called a cyclic operator [16].

The following corollary is the fuzzy metric version of the result of Kirk et al. [16].

Corollary 3. *Let $(X, M, *)$ be a complete fuzzy metric space, m be a positive integer, $A_i, i = 1, 2, \dots, m$ be nonempty closed subsets of X and $B = \bigcup_{i=1}^m A_i$ be a cyclic representation of B with respect to T. Suppose $A_{m+i} = A_i$ for all $i \in \mathbb{N}$ and following condition holds:*

$$\exists_{\lambda\in(0,1)}\ \Bigg(x \in A_i\,,\ y \in A_{i+1}\,,\ i = 1, 2, \dots, m \\ \Rightarrow\ \frac{1}{M(Tx,Ty,t)} - 1 \le k\left[\frac{1}{M(x,y,t)} - 1\right]\Bigg).$$

Then T has a unique fixed point $x^* \in \bigcap_{i=1}^m A_i$.

Proof. Since $B = \bigcup_{i=1}^{m} A_i$ is closed so $(B, M, *)$ is complete. Let G be the graph defined by $V(G) = B$ and

$$E(G) = \Delta \cup \big\{(x, y) \in B \times B : x \in A_i \wedge y \in A_{i+1} \colon i = 1, 2, \ldots, m\big\}.$$

Then since $B = \bigcup_{i=1}^{m} A_i$ is a cyclic representation of B with respect to T, so (GF1) holds and by the given contractive condition (GF2) also hold. Now it is easy to see that the sequence $(T^n x)_{n\in\mathbb{N}}$ has infinitely many terms in each $A_i, i = 1, 2, \ldots, m$ so if $(T^n x)_{n\in\mathbb{N}}$ converges to x^* then $x^* \in \bigcap_{i=1}^{m} A_i$. Therefore $(\mathcal{P}_T)$ holds good. Note that if $x \in B$ then $(x, Tx) \in E(G)$ therefore $T \subseteq E(G)$ and by (D) of Theorem 3, T has a fixed point. Uniqueness follows from the contractive condition and the fact that if $x \in \operatorname{Fix} T$ then $x \in \bigcap_{i=1}^{m} A_i$. □

References

[1] A. George, P. Veeramani: On some results in fuzzy metric spaces. *Fuzzy Sets and Systems* 64 (1994) 395–399.

[2] A. George, P. Veeramani: On some results of analysis for fuzzy metric spaces. *Fuzzy Sets Systems* 90 (1997) 365–368.

[3] A. Petruşel, I.A. Rus: Fixed point theorems in ordered L-spaces. *Proc. Amer. Math. Soc.* 134 (2006) 411–418. MR2176009 (2006g:47097)

[4] B. Schweizer, A. Sklar: Statistical metric spaces. *Pacific J. Math.* 10 (1960) 313–334.

[5] E.P. Klement, R. Mesiar, E. Pap: *Triangular Norms, Trends in Logics, vol. 8.* Kluwer Academic Publishers, Dordrecht, Boston, London (2000).

[6] F. Bojor: Fixed points of Kannan mappings in metric spaces endowed with a graph. *An. Şt. Univ. Ovidius Constanţa* 20 (1) (2012) 31–40. DOI: 10.2478/v10309-012-0003-x

[7] I. Kramosil, J. Michálek: Fuzzy metrics and statistical metric spaces. *Kybernetika* 11 (1975) 336–344.

[8] J.J. Nieto, R.L. Pouso, R. Rodríguez-López: Fixed point theorems in ordered abstract spaces. *Proc. Amer. Math. Soc.* 135 (2007) 2505—2517. MR2302571

[9] J.J. Nieto, R. Rodríguez–López: Contractive Mapping Theorems in Partially Ordered Sets and Applications to Ordinary Differential Equations. *Order* 22 (2005) 223–239.

[10] J.J. Nieto, R. Rodríguez-López: Existence and uniqueness of fixed point in partially ordered sets and applications to ordinary differential equations. *Acta Math. Sinica, English Ser.* (2007) 2205–2212.

[11] J. Jachymski: The contraction principle for mappings on a metric space with a graph. *Proc. Amer. Math. Soc.* 136 (2008) 1359–1373.

[12] L.A. Zadeh: Fuzzy sets. *Information and Control* 89 (1965) 338–353.

[13] M. Grabiec: Fixed points in fuzzy metric spaces. *Fuzzy Sets and Systems* 27 (1988) 385–389.

[14] O. Hadžić, E. Pap: *Fixed Point Theory in Probabilistic Metric Spaces.* Mathematics and its Applications. Vol. 536. Kluwer Academic Publishers, Dordrecht, Boston, London (2001).

[15] V. Gregori, A. Sapena: On fixed-point theorems in fuzzy metric spaces. *Fuzzy Sets and Systems* 125 (2002) 245–252.

[16] W.A. Kirk, P.S. Srinavasan, P. Veeramani: Fixed points for mappings satisfying cyclical contractive conditions. *Fixed Point Theory* 4 (2003) 79–89.

Author's address:
Department of Applied Mathematics, Shri Vaishnav Institute of Technology & Science, Sanwer Road, Indore, (M.P.) 453331, India
E-mail: satishmathematics@yahoo.co.in

Received: 1 September, 2013
Accepted for publication: 10 March, 2014
Communicated by: Vilém Novák

Communications in Mathematics 22 (2014) 13–30

New hyper-Kähler structures on tangent bundles

Xuerong Qi, Linfen Cao, Xingxiao Li

Abstract. Let (M, g, J) be an almost Hermitian manifold, then the tangent bundle TM carries a class of naturally defined almost hyper-Hermitian structures (G, J_1, J_2, J_3). In this paper we give conditions under which these almost hyper-Hermitian structures (G, J_1, J_2, J_3) are locally conformal hyper-Kähler. As an application, a family of new hyper-Kähler structures is obtained on the tangent bundle of a complex space form. Furthermore, by restricting these almost hyper-Hermitian structures on the unit tangent sphere bundle T_1M, we obtain a class of almost contact metric 3-structures. By virtue of these almost contact metric 3-structures, we find a family of Sasakian 3-structures on the unit tangent sphere bundle of a complex space form of positive holomorphic sectional curvature.

1 Introduction

A Riemannian metric g on a smooth manifold M gives rise to several natural Riemannian metrics and almost complex structures on the tangent bundle TM and the cotangent bundle T^*M of M. Maybe the best known examples are the Sasaki metric g_s and the canonical almost complex structure J_s (see [11], [18]). The Sasaki metric g_s and the canonical almost complex structure J_s determine an almost Hermitian structure on TM. Although the natural almost Hermitian structure (g_s, J_s) is almost Kähler, it is very rigid in the following sense. For example, the Sasaki metric g_s has never constant scalar curvature unless the metric g on the base manifold M is flat (see [7], [10]). On the other hand, the canonical almost complex structure J_s is integrable if and only if the base manifold (M, g) is flat (see [6], [19]). The rigidity of the natural almost Hermitian structure (g_s, J_s) has incited many authors to tackle the problem of the construction and the study of other almost Hermitian structures on TM or T^*M ([1], [9], [12], [14], [17], [22]). Especially, Oproiu and Papaghiuc [15] has constructed a class of Kähler-Einstein structures on the cotangent bundle of a real space form. Recently, the authors [8]

2010 MSC: 53C15, 53C26
Key words: tangent bundles, locally conformal hyper-Kähler structures, almost contact metric 3-structures, Sasakian 3-structures

have proved that the tangent bundle of any Riemannian manifold admits a class of locally conformal almost Kähler structures. By virtue of these structures, they have also shown that there exists a class of Sasakian structures on the unit tangent sphere bundle of a real space form of positive sectional curvature.

For an almost Hermitian manifold (M, g, J), almost hyper-Hermitian structures can be found on TM and T^*M ([3], [4], [20]). In particular, Calabi [4] constructed a hyper-Kähler structure on the cotangent bundle of a complex projective space $\mathbb{C}P^n$. Tahara, Vanhecke and Watanabe [20] gave a class of almost hyper-Hermitian structures (G, J_1, J_2, J_3) on TM, which is determined by some parameters. Furthermore, suitably choosing these parameters, they obtained a family of hyper-Kähler structures on the tangent bundle of a complex space form of positive holomorphic sectional curvature. Oproiu [13] studied a family of almost hyper-Hermitian structures on the tangent bundle of a Kähler manifold, and obtained the necessary and sufficient conditions for these almost hyper-Hermitian structures to be hyper-Kähler structures.

In this paper, we study the geometry of these almost hyper-Hermitian structures (G, J_1, J_2, J_3) on the tangent bundle TM of an almost Hermitian manifold (M, g, J). The arrangement of this paper is as follows: section 2 are some necessary preliminaries and known results. In section 3, we give conditions under which these almost hyper-Hermitian structures (G, J_1, J_2, J_3) are locally conformal hyper-Kähler. As an application, a class of new hyper-Kähler structures is obtained on the tangent bundle of a complex space form. Our result is more general than one of [20]. It shows that there also exists a class of hyper-Kähler structures on the tangent bundle of a complex space form of non-positive holomorphic sectional curvature. In the end of this section, we present some concrete examples of hyper-Kähler structures on the tangent bundle of a complex space form. In section 4, using these almost hyper-Hermitian structures (G, J_1, J_2, J_3), we obtain a class of almost contact metric 3-structures $(\varphi_\alpha, \xi_\alpha, \eta_\alpha, \widehat{G})_{\alpha=1,2,3}$ on the unit tangent sphere bundle T_1M. By studying some properties of these almost contact metric 3-structures $(\varphi_\alpha, \xi_\alpha, \eta_\alpha, \widehat{G})_{\alpha=1,2,3}$, we find a family of Sasakian 3-structures on the unit tangent sphere bundle of a complex space form of positive holomorphic sectional curvature.

To simplify matters, we first make the following conventions on the ranges of indices used frequently in this paper:

$$1 \le i, j, k, \ldots \le n, \quad 1 \le \alpha, \beta, \gamma, \ldots \le 3\,.$$

2 Preliminaries

Let (M, g) be a smooth n-dimensional Riemannian manifold and denote its tangent bundle by $\pi\colon TM \to M$. The Levi-Civita connection ∇ of g defines a direct sum decomposition

$$TTM = HTM \oplus VTM \tag{1}$$

of the tangent bundle TTM into the vertical distribution $VTM = \ker \pi_*$ and the horizontal distribution HTM. Locally, if $(U; x^1, \ldots, x^n)$ is a local coordinate system on M, then $(\pi^{-1}(U); x^1, \ldots, x^n, y^1, \ldots, y^n)$ is a local coordinate system on TM. The metric g can be locally expressed as

$$g = g_{ij}(x)dx^i dx^j, \quad x \in U.$$

Denote by Γ^k_{ij} the Christoffel symbols of g and define

$$\Gamma^k_{i0}(x,y) = \Gamma^k_{ij}(x)y^j, \quad \text{for } y = y^i \frac{\partial}{\partial x^i}(x), \ x \in U,$$

then

$$\left\{\delta_i := \frac{\partial}{\partial x^i} - \Gamma^k_{i0}\frac{\partial}{\partial y^k}\right\}, \quad \left\{\partial_i := \frac{\partial}{\partial y^i}\right\} \tag{2}$$

define respectively a local frame field for HTM and a local frame field for VTM over $\pi^{-1}(U)$. Therefore $\{\delta_1, \ldots, \delta_n, \partial_1, \ldots, \partial_n\}$ defines a local tangent frame field on TM, adapted to the direct sum decomposition (1). Put

$$\nabla y^i = dy^i + \Gamma^i_{0k}dx^k,$$

then $\{dx^1, \ldots, dx^n, \nabla y^1, \ldots, \nabla y^n\}$ is a local cotangent frame field on TM dual to the local frame field $\{\delta_1, \ldots, \delta_n, \partial_1, \ldots, \partial_n\}$.

If we denote respectively by X^H and X^V the horizontal and the vertical lift to TM of a tangent vector field X on M, then

$$\delta_i = \left(\frac{\partial}{\partial x^i}\right)^H, \quad \partial_i = \left(\frac{\partial}{\partial x^i}\right)^V. \tag{3}$$

For any $x \in U$ and $y = y^i \frac{\partial}{\partial x^i}(x) \in \pi^{-1}(x)$, or in other words, for $(x,y) \in \pi^{-1}(U)$, put

$$t = \frac{1}{2}g_{\pi(y)}(y,y) = \frac{1}{2}g_{ij}(x)y^i y^j,$$

then it easily follows that

$$\delta_i t = 0, \quad \partial_i t = g_{ji}y^j := g_{0i}. \tag{4}$$

Furthermore, we define

$$R^l_{0ij}(x,y) = y^k R^l_{kij}(x), \quad \text{for } (x,y) \in \pi^{-1}(U),$$

then the Lie brackets of the vector fields ∂_i, δ_i are given by

$$[\partial_i, \partial_j] = 0, \qquad [\delta_i, \partial_j] = \Gamma^l_{ij}\partial_l, \qquad [\delta_i, \delta_j] = -R^l_{0ij}\partial_l. \tag{5}$$

The Sasaki metric g_s is uniquely determined by the following equations:

$$g_s(X^H, Y^H) = g_s(X^V, Y^V) = g(X,Y) \circ \pi,$$
$$g_s(X^H, Y^V) = 0, \quad \forall X, Y \in \mathscr{X}(M),$$

where $\mathscr{X}(M)$ denotes the Lie algebra of smooth tangent vector fields on M. The canonical almost complex structure J_s on TM is given by

$$J_s X^H = X^V, \quad J_s X^V = -X^H, \quad \forall X \in \mathscr{X}(M).$$

It is known that (TM, g_s, J_s) is an almost Kähler manifold. Moreover, the integrability of the almost complex structure J_s implies that (M, g) is locally Euclidean (see [6], [19]).

The Cheeger-Gromoll metric g_{CG} [5] is uniquely determined by

$$\begin{cases} g_{CG}(X^H, Y^H) = g(X,Y)\circ\pi\,, \\ g_{CG}(X^V, Y^V) = \dfrac{1}{1+2t}\left[g(X,Y) + g(X,y)g(Y,y)\right], \\ g_{CG}(X^H, Y^V) = 0\,, \end{cases}$$

where $X, Y \in \mathscr{X}(M)$ and $y \in TM$. Accordingly, one can define an almost complex structure J_{CG} on TM by the following equations:

$$\begin{aligned} J_{CG}X^H &= \tau X^V - \frac{1}{1+\tau} g(X,y)y^V, \\ J_{CG}X^V &= -\frac{1}{\tau}X^H - \frac{1}{\tau(1+\tau)} g(X,y)y^H, \end{aligned}$$

where $\tau = \sqrt{1+2t}$. Then (TM, g_{CG}, J_{CG}) is an almost Hermitian manifold (see [9]).

As a generalization of the above two metrics, one can define a class of metrics G on the tangent bundle TM of an almost Hermitian manifold (M, g, J) by the following equations (see [20]):

$$\begin{cases} G(X^H, Y^H) = c_1 g(X,Y) + d_1 g(X,y)g(Y,y) + f_1 g(X,Jy)g(Y,Jy)\,, \\ G(X^V, Y^V) = c_2 g(X,Y) + d_2 g(X,y)g(Y,y) + f_2 g(X,Jy)g(Y,Jy)\,, \\ G(X^H, Y^V) = 0\,, \end{cases}$$

where $c_1, c_2, d_1, d_2, f_1, f_2$ are smooth functions of $t \in [0, \infty)$, and satisfy the following conditions:

$$\begin{gathered} c_1 > 0\,, \quad c_2 > 0\,, \quad c_1 + 2td_1 > 0\,, \\ c_2 + 2td_2 > 0\,, \quad c_1 + 2tf_1 > 0, \quad c_2 + 2tf_2 > 0\,, \qquad \forall t\,. \end{gathered}$$

Accordingly, one can also define three kinds of almost complex structures J_1, J_2, J_3 on TM by the following equations (see [20]):

$$\begin{cases} J_1 X^H = a_1 X^V + b_1 g(X,y)y^V + e_1 g(X,Jy)(Jy)^V, \\ J_1 X^V = -a_2 X^H - b_2 g(X,y)y^H - e_2 g(X,Jy)(Jy)^H, \end{cases}$$

$$\begin{cases} J_2 X^H = a_1 (JX)^V + b_1 g(JX,y)y^V + e_1 g(X,y)(Jy)^V, \\ J_2 X^V = a_2 (JX)^H + e_2 g(JX,y)y^H + b_2 g(X,y)(Jy)^H, \end{cases}$$

$$\begin{cases} J_3 X^H = -(JX)^H, \\ J_3 X^V = (JX)^V + pg(JX,y)y^V + qg(X,y)(Jy)^V, \end{cases}$$

where $p = a_2 b_1 + a_1 e_2 + 2t b_1 e_2$, $q = a_2 e_1 + a_1 b_2 + 2t b_2 e_1$, a_1, a_2, b_1, b_2, e_1, e_2 are smooth functions of t satisfying the following conditions:

$$a_1 a_2 = 1\,, \quad (a_1 + 2tb_1)(a_2 + 2tb_2) = 1\,, \quad (a_1 + 2te_1)(a_2 + 2te_2) = 1\,, \tag{6}$$

or

$$a_2 = \frac{1}{a_1}\,, \quad b_2 = -\frac{b_1}{a_1(a_1 + 2tb_1)}\,, \quad e_2 = -\frac{e_1}{a_1(a_1 + 2te_1)}\,. \tag{7}$$

A direct computation shows that $J_3 = J_1 \circ J_2 = -J_2 \circ J_1$, so (J_1, J_2, J_3) is an almost hyper-complex structure on TM. Moreover, (7) shows that the almost hyper--complex structure (J_1, J_2, J_3) depends on three essential parameters a_1, b_1, e_1.

Remark 1. From (6) we know that the coefficients a_1, a_2, $a_1 + 2tb_1$, $a_2 + 2tb_2$, $a_1 + 2te_1$, $a_2 + 2te_2$ cannot vanish and have the same sign. In this paper, we assume that

$$\begin{gathered} a_1 > 0\,, \quad a_2 > 0\,, \quad a_1 + 2tb_1 > 0\,, \\ a_2 + 2tb_2 > 0\,, \quad a_1 + 2te_1 > 0\,, \quad a_2 + 2te_2 > 0\,, \qquad \forall t\,. \end{gathered}$$

For the metric G and almost hyper-complex structure (J_1, J_2, J_3), we have the following results. These results follow from corresponding statements in [20].

Proposition 1. *The almost complex structures J_1, J_2, J_3 are compatible with the metric G if and only if*

$$\begin{aligned} &\frac{c_1}{a_1} = \frac{c_2}{a_2} = \lambda\,, && \frac{c_1 + 2td_1}{a_1 + 2tb_1} = \frac{c_2 + 2td_2}{a_2 + 2tb_2} = \lambda + 2t\mu\,, \\ &d_1 = f_1\,, && \frac{c_1 + 2tf_1}{a_1 + 2te_1} = \frac{c_2 + 2tf_2}{a_2 + 2te_2} = \lambda + 2t\nu\,, \end{aligned}$$

where the proportionality coefficients $\lambda > 0$, $\lambda + 2t\mu > 0$ and $\lambda + 2t\nu > 0$ for all t.

Remark 2. Proposition 1 shows that the almost hyper-Hermitian structure (G, J_1, J_2, J_3) depends on five essential parameters a_1, b_1, e_1, λ, μ or a_1, b_1, e_1, λ, ν. Other parameters are determined by

$$a_2 = \frac{1}{a_1}\,, \quad b_2 = -\frac{b_1}{a_1(a_1 + 2tb_1)}\,, \quad e_2 = -\frac{e_1}{a_1(a_1 + 2te_1)}\,, \tag{8}$$

$$c_1 = \lambda a_1\,, \quad d_1 = f_1 = \lambda b_1 + \mu(a_1 + 2tb_1) = \lambda e_1 + \nu(a_1 + 2te_1)\,, \tag{9}$$

$$c_2 = \lambda a_2\,, \quad d_2 = \lambda b_2 + \mu(a_2 + 2tb_2)\,, \quad f_2 = \lambda e_2 + \nu(a_2 + 2te_2)\,. \tag{10}$$

Proposition 2. *Let (M, g, J) be a complex space form of holomorphic sectional curvature $4c$. If the functions b_1, e_1 are given by*

$$b_1 = \frac{a_1 a_1' - c}{a_1 - 2ta_1'}\,, \quad e_1 = \frac{c}{a_1}\,, \tag{11}$$

then the almost hyper-complex structures (J_1, J_2, J_3) is a hyper-complex structure on TM.

Remark 3. If we choose

$$a_1 = a_2 = c_1 = c_2 = 1\,, \quad b_1 = b_2 = e_1 = e_2 = d_1 = d_2 = f_1 = f_2 = 0\,,$$

then (G, J_1, J_2, J_3) is the canonical almost hyper-Hermitian structure on the tangent bundle TM. In this case, $G = g_s$, $J_1 = J_s$. Therefore, the canonical almost hyper-Hermitian structure (G, J_1, J_2, J_3) is a hyper-Hermitian structure if and only if the base manifold (M, g) if flat.

3 The tangent bundle TM

In the sequel (M, g, J) will be a connected almost Hermitian manifold. We study the geometry of the almost hyper-Hermitian manifold $(TM, G, J_\alpha)_{\alpha=1,2,3}$ and find conditions under which the considered almost hyper-Hermitian structure is locally conformal hyper-Kähler or hyper-Kähler. Now we first introduce the definitions of locally conformal Kähler manifold and locally conformal hyper-Kähler manifold (see also [16], [21]).

Definition 1. The almost Hermitian manifold $(\widetilde{M}, \widetilde{g}, \widetilde{J})$ is called a locally conformal almost Kähler manifold if there exists a closed 1-form $\widetilde{\omega}$, called the *Lee form*, satisfying $d\widetilde{\Omega} = \widetilde{\omega} \wedge \widetilde{\Omega}$, where $\widetilde{\Omega}$ is the fundamental 2-form. Moreover, the locally conformal almost Kähler manifold $(\widetilde{M}, \widetilde{g}, \widetilde{J})$ is called a locally conformal Kähler manifold if $\widetilde{J}$ is integrable.

Definition 2. The hyper-Hermitian manifold $(\widetilde{M}, \widetilde{g}, \widetilde{J}_\alpha)_{\alpha=1,2,3}$ is called a locally conformal hyper-Kähler manifold if there exists a closed 1-form $\widetilde{\omega}$, such that, for each α, $d\widetilde{\Omega}_\alpha = \widetilde{\omega} \wedge \widetilde{\Omega}_\alpha$, where $\widetilde{\Omega}_\alpha$ is the fundamental 2-form of $(\widetilde{g}, \widetilde{J}_\alpha)$.

Denote by Ω_α the fundamental 2-form of (G, J_α), given by $\Omega_\alpha(\cdot,\cdot) = G(\cdot, J_\alpha \cdot)$, then under the adapted frame field $\{\delta_1, \ldots, \delta_n, \partial_1, \ldots, \partial_n\}$ defined by (2) or (3), we have

$$\Omega_1 = (\lambda g_{ij} + \mu g_{i0} g_{j0} + \nu g_{i\bar{0}} g_{j\bar{0}}) \nabla y^i \wedge dx^j, \tag{12}$$

$$\Omega_2 = (\lambda g_{i\bar{j}} + \mu g_{i0} g_{\bar{j}0} + \nu g_{i\bar{0}} g_{j0}) \nabla y^i \wedge dx^j, \tag{13}$$

$$\begin{aligned} \Omega_3 = \{ & c_2 g_{i\bar{j}} + [\lambda e_2 + \mu(a_2 + 2te_2)] g_{i0} g_{\bar{j}0} \\ & + [\lambda b_2 + \nu(a_2 + 2tb_2)] g_{i\bar{0}} g_{j0} \} \nabla y^i \wedge \nabla y^j \\ & - (c_1 g_{i\bar{j}} + d_1 g_{i0} g_{\bar{j}0} + f_1 g_{i\bar{0}} g_{j0}) dx^i \wedge dx^j, \end{aligned} \tag{14}$$

where $g_{i\bar{j}} = g\big(\frac{\partial}{\partial x^i}, J\frac{\partial}{\partial x^j}\big)$, $g_{i\bar{0}} = g\big(\frac{\partial}{\partial x^i}, Jy\big)$.

By using (4) and the property that $\nabla g = 0$, we obtain

$$d\lambda = \lambda' g_{i0} \nabla y^i, \qquad d\mu = \mu' g_{i0} \nabla y^i, \qquad d\nu = \nu' g_{i0} \nabla y^i,$$

$$dg_{i0} = g_{ik} \nabla y^k + \Gamma^l_{ik} g_{l0} dx^k, \qquad dg_{\bar{i}0} = -dg_{i\bar{0}} = g_{\bar{i}k} \nabla y^k + \Gamma^l_{\bar{i}k} g_{l0} dx^k,$$

$$d\nabla y^i = \Gamma^i_{lk} \nabla y^l \wedge dx^k + \frac{1}{2} R^i_{0kl} dx^k \wedge dx^l,$$

where $\Gamma^l_{\bar{i}k} g_{l0} = g\big(\nabla_{\frac{\partial}{\partial x^k}} J\frac{\partial}{\partial x^i}, y\big)$.

Now we consider the first class of almost Hermitian structures (G, J_1). By the local expression (12) of Ω_1, we have

$$
\begin{aligned}
d\Omega_1 &= d(\lambda g_{ij} + \mu g_{i0} g_{j0} + \nu g_{i\bar{0}} g_{j\bar{0}}) \wedge \nabla y^i \wedge dx^j \\
&\quad + (\lambda g_{ij} + \mu g_{i0} g_{j0} + \nu g_{i\bar{0}} g_{j\bar{0}}) d\nabla y^i \wedge dx^j \\
&= \big\{\lambda' g_{ij} g_{0k} \nabla y^k + \lambda(\Gamma^l_{ik} g_{lj} + \Gamma^l_{jk} g_{li}) dx^k + \mu' g_{0i} g_{0j} g_{0k} \nabla y^k \\
&\quad + \mu g_{0j}(g_{ik} \nabla y^k + \Gamma^l_{ik} g_{0l} dx^k) + \mu g_{0i}(g_{jk} \nabla y^k + \Gamma^l_{jk} g_{0l} dx^k) + \nu' g_{i\bar{0}} g_{j\bar{0}} g_{0k} \nabla y^k \\
&\quad + \nu g_{j\bar{0}}(g_{i\bar{k}} \nabla y^k - \Gamma^l_{\bar{i}k} g_{l0} dx^k) + \nu g_{i\bar{0}}(g_{j\bar{k}} \nabla y^k - \Gamma^l_{\bar{j}k} g_{l0} dx^k)\big\} \wedge \nabla y^i \wedge dx^j \\
&\quad + (\lambda g_{ij} + \mu g_{0i} g_{0j} + \nu g_{i\bar{0}} g_{j\bar{0}}) \Big(\Gamma^i_{lk} \nabla y^l \wedge dx^k + \frac{1}{2} R^i_{0kl} dx^k \wedge dx^l\Big) \wedge dx^j .
\end{aligned}
$$

Clearly, if (TM, G, J_1) is a locally conformal almost Kähler manifold, then

$$
\frac{1}{2}(\lambda g_{ij} + \mu g_{0i} g_{0j} + \nu g_{i\bar{0}} g_{j\bar{0}}) R^i_{0kl} dx^k \wedge dx^l \wedge dx^j = 0 . \tag{15}
$$

Using the first Bianchi identity, (15) is equivalent to

$$
\nu g_{j\bar{0}} R_{0\bar{0}kl} dx^k \wedge dx^l \wedge dx^j = 0 ,
$$

where $R_{0\bar{0}kl} = R(y, Jy, \frac{\partial}{\partial x^k}, \frac{\partial}{\partial x^l})$. Let $\nu = 0$, then we find that

$$
\Omega_1 = (\lambda g_{ij} + \mu g_{i0} g_{j0}) \nabla y^i \wedge dx^j , \tag{16}
$$

$$
d\Omega_1 = (\lambda' - \mu) g_{ij} g_{0k} \nabla y^k \wedge \nabla y^i \wedge dx^j . \tag{17}
$$

Therefore, if $\nu = 0$ and $\mu = \lambda'$, then (TM, G, J_1) is almost Kähler.

Proposition 3. *If $\nu = 0$, then the almost Hermitian manifold (TM, G, J_1) is locally conformal almost Kähler.*

Proof. Set

$$
\omega = \frac{\lambda' - \mu}{\lambda} g_{0k} \nabla y^k . \tag{18}
$$

Then ω is a globally defined 1-form. Moreover

$$
\begin{aligned}
d\omega &= \left(\frac{\lambda' - \mu}{\lambda} g_{0i}\right) d\nabla y^i + \left(\frac{\lambda' - \mu}{\lambda}\right)' g_{0i} g_{0k} \nabla y^i \wedge \nabla y^k \\
&\quad + \frac{\lambda' - \mu}{\lambda} \left(g_{ik} \nabla y^k + \Gamma^l_{ik} g_{0l} dx^k\right) \wedge \nabla y^i \\
&= \frac{\lambda' - \mu}{\lambda} g_{0i} \left(\Gamma^i_{lk} \nabla y^l \wedge dx^k + \frac{1}{2} R^i_{0kl} dx^k \wedge dx^l\right) \\
&\quad + \frac{\lambda' - \mu}{\lambda} \Gamma^l_{ik} g_{0l} dx^k \wedge \nabla y^i = 0 .
\end{aligned}
$$

On the other hand, it follows from (17) that

$$\begin{aligned} d\Omega_1 &= (\lambda' - \mu) g_{ij} g_{0k} \nabla y^k \wedge \nabla y^i \wedge dx^j \\ &= \left(\frac{\lambda' - \mu}{\lambda} g_{0k} \nabla y^k \right) \wedge (\lambda g_{ij} + \mu g_{0i} g_{0j}) \nabla y^i \wedge dx^j \\ &\quad - \left(\frac{\mu(\lambda' - \mu)}{\lambda} g_{0k} g_{0i} g_{0j} \right) \nabla y^k \wedge \nabla y^i \wedge dx^j . \end{aligned} \tag{19}$$

The second term in (19) vanishes since $g_{0i} g_{0k}$ is symmetric with respect to the indices i, k. Hence

$$d\Omega_1 = \omega \wedge \Omega_1. \tag{20}$$

The closeness of ω and (20) show that (TM, G, J_1) is a locally conformal almost Kähler manifold. □

The following corollary comes directly from Proposition 2 and Proposition 3.

Corollary 1. *Let (M, g, J) be a complex space form of holomorphic sectional curvature $4c$. If the functions b_1, e_1 are given by (11) and $\nu = 0$, then the almost Hermitian manifold (TM, G, J_1) is a locally conformal Kähler manifold. Moreover, the locally conformal Kähler manifold (TM, G, J_1) is a Kähler manifold if $\mu = \lambda'$.*

Next we consider the second class of almost Hermitian structures (G, J_2). By the local expression (13) of Ω_2, we have

$$\begin{aligned} d\Omega_2 &= d(\lambda g_{i\bar{j}} + \mu g_{i0} g_{\bar{j}0} + \nu_{i\bar{0}} g_{j0}) \wedge \nabla y^i \wedge dx^j \\ &\quad + (\lambda g_{i\bar{j}} + \mu g_{i0} g_{\bar{j}0} + \nu_{i\bar{0}} g_{j0}) d\nabla y^i \wedge dx^j \\ &= \big\{ \lambda' g_{i\bar{j}} g_{0k} \nabla y^k + \lambda (\Gamma^l_{ik} g_{l\bar{j}} + \Gamma^l_{\bar{j}k} g_{li}) dx^k + \mu' g_{0i} g_{\bar{j}0} g_{0k} \nabla y^k \\ &\quad + \mu g_{0\bar{j}} (g_{ik} \nabla y^k + \Gamma^l_{ik} g_{0l} dx^k) + \mu g_{0i} (g_{\bar{j}k} \nabla y^k + \Gamma^l_{\bar{j}k} g_{0l} dx^k) + \nu' g_{i\bar{0}} g_{j0} g_{0k} \nabla y^k \\ &\quad + \nu g_{j0} (g_{i\bar{k}} \nabla y^k - \Gamma^l_{\bar{i}k} g_{l0} dx^k) + \nu g_{i\bar{0}} (g_{jk} \nabla y^k + \Gamma^l_{jk} g_{l0} dx^k) \big\} \wedge \nabla y^i \wedge dx^j \\ &\quad + (\lambda g_{i\bar{j}} + \mu g_{0i} g_{\bar{j}0} + \nu g_{i\bar{0}} g_{j0}) \Big(\Gamma^i_{lk} \nabla y^l \wedge dx^k + \frac{1}{2} R^i_{0kl} dx^k \wedge dx^l \Big) \wedge dx^j . \end{aligned}$$

Clearly, if (TM, G, J_2) is a locally conformal almost Kähler manifold, then

$$\frac{1}{2} (\lambda g_{i\bar{j}} + \mu g_{0i} g_{\bar{j}0} + \nu g_{i\bar{0}} g_{j0}) R^i_{0kl} dx^k \wedge dx^l \wedge dx^j = 0 \,. \tag{21}$$

In the case that (M, g, J) is a Kähler manifold, using the first Bianchi identity, (21) is equivalent to

$$\nu g_{j0} R_{0\bar{0}kl} dx^k \wedge dx^l \wedge dx^j = 0.$$

Let $\nu = 0$, then we obtain that

$$\Omega_2 = (\lambda g_{i\bar{j}} + \mu g_{i0} g_{\bar{j}0}) \nabla y^i \wedge dx^j , \tag{22}$$

$$d\Omega_2 = (\lambda' - \mu) g_{i\bar{j}} g_{0k} \nabla y^k \wedge \nabla y^i \wedge dx^j . \tag{23}$$

Proposition 4. *Let (M,g,J) be a Kähler manifold. If $\nu=0$, then the almost Hermitian manifold (TM,G,J_2) is locally conformal almost Kähler.*

Proof. From (18), (22) and (23), it follows that

$$\begin{aligned} d\Omega_2 &= (\lambda'-\mu)g_{i\bar{j}}g_{0k}\nabla y^k\wedge\nabla y^i\wedge dx^j \\ &= \left(\frac{\lambda'-\mu}{\lambda}g_{0k}\nabla y^k\right)\wedge(\lambda g_{i\bar{j}}+\mu g_{0i}g_{0\bar{j}})\nabla y^i\wedge dx^j \\ &= \omega\wedge\Omega_2\,. \end{aligned} \tag{24}$$

The closeness of ω and (24) show that (TM,G,J_2) is a locally conformal almost Kähler manifold. □

The following corollary comes directly from Proposition 2 and Proposition 4:

Corollary 2. *Let (M,g,J) be a complex space form of holomorphic sectional curvature $4c$. If the functions b_1,e_1 are given by (11) and $\nu=0$, then the almost Hermitian manifold (TM,G,J_2) is a locally conformal Kähler manifold. Moreover, the locally conformal Kähler manifold (TM,G,J_2) is a Kähler manifold if $\mu=\lambda'$.*

For the third class of almost Hermitian structures (G,J_3), we have

$$\begin{aligned} d\Omega_3 = &\big(-c_1'g_{i\bar{j}}g_{0k}-d_1'g_{0i}g_{0\bar{j}}g_{0k}-d_1'g_{\bar{0}i}g_{0j}g_{0k}-d_1g_{0\bar{j}}g_{ik}-d_1g_{\bar{0}i}g_{jk} \\ &-d_1g_{0i}g_{\bar{j}k}-d_1g_{0j}g_{i\bar{k}}+c_2R_{0\bar{k}ij}+\frac{\sigma}{2}g_{0k}R_{0\bar{0}ij}\big)\nabla y^k\wedge dx^i\wedge dx^j \\ &+(-c_1\Gamma^l_{\bar{j}k}g_{li}-2d_1g_{0i}\Gamma^l_{\bar{j}k}g_{l0})dx^k\wedge dx^i\wedge dx^j \\ &+(c_2'g_{i\bar{j}}g_{0k}+\sigma g_{0i}g_{\bar{j}k})\nabla y^k\wedge\nabla y^i\wedge\nabla y^j \\ &+(c_2\Gamma^l_{\bar{j}k}g_{li}-c_2\Gamma^l_{jk}g_{\bar{l}i}+\sigma g_{0i}\Gamma^l_{\bar{j}k}g_{l0}-\sigma g_{0i}\Gamma^l_{jk}g_{\bar{l}0})dx^k\wedge\nabla y^i\wedge\nabla y^j\,, \end{aligned}$$

where $\sigma=\lambda(b_2+e_2)+\mu(a_2+2te_2)+\nu(a_2+2tb_2)$.

In the case that (M,g,J) is a complex space form of holomorphic sectional curvature $4c$, we find

$$\begin{aligned} R_{0\bar{k}ij} &= -c(g_{0i}g_{\bar{k}j}-g_{0j}g_{\bar{k}i}+g_{0\bar{i}}g_{kj}-g_{0\bar{j}}g_{ki}-2g_{0k}g_{i\bar{j}})\,, \\ R_{0\bar{0}ij} &= -2c(g_{0i}g_{\bar{0}j}+g_{0\bar{i}}g_{0j}-2tg_{i\bar{j}})\,. \end{aligned}$$

Therefore,

$$\begin{aligned} c_2R_{0\bar{k}ij}+\frac{\sigma}{2}g_{0k}R_{0\bar{0}ij} = &-c\sigma g_{0k}(g_{0i}g_{\bar{0}j}+g_{0\bar{i}}g_{0j}-2tg_{i\bar{j}}) \\ &-cc_2(g_{0i}g_{\bar{k}j}-g_{0j}g_{\bar{k}i}+g_{0\bar{i}}g_{kj}-g_{0\bar{j}}g_{ki}-2g_{0k}g_{i\bar{j}})\,. \end{aligned}$$

Consequently,

$$\begin{aligned} d\Omega_3 = &(c_2'-\sigma)g_{i\bar{j}}g_{0k}\nabla y^k\wedge\nabla y^i\wedge\nabla y^j+\big\{-c_1'g_{i\bar{j}}g_{0k}-d_1'g_{0i}g_{0\bar{j}}g_{0k} \\ &-d_1'g_{\bar{0}i}g_{0j}g_{0k}-(d_1-cc_2)(g_{0\bar{j}}g_{ik}+g_{\bar{0}i}g_{jk}+g_{0i}g_{\bar{j}k}+g_{0j}g_{i\bar{k}}) \\ &+2cc_2g_{0k}g_{i\bar{j}}-c\sigma g_{0k}(g_{0i}g_{\bar{0}j}+g_{0\bar{i}}g_{0j}-2tg_{i\bar{j}})\big\}\nabla y^k\wedge dx^i\wedge dx^j\,. \end{aligned} \tag{25}$$

Theorem 1. *Let (M, g, J) be a complex space form of holomorphic sectional curvature $4c$. If the functions b_1, e_1, ν are given by*

$$b_1 = \frac{a_1 a_1' - c}{a_1 - 2ta_1'}\,, \qquad e_1 = \frac{c}{a_1}\,, \qquad \nu = 0\,, \tag{26}$$

then the almost Hermitian structure (G, J_3) is locally conformal Kähler. In this case, the almost hyper-Hermitian structure (G, J_1, J_2, J_3) is locally conformal hyper-Kähler.

Proof. Set $\varrho = \frac{\lambda' - \mu}{\lambda}$. Then $\omega = \varrho g_{0k} \nabla y^k$ is a globally defined 1-form and $d\omega = 0$. From (14), it follows easily that

$$\begin{aligned} \omega \wedge \Omega_3 = \big\{ & \varrho c_2 g_{i\bar{j}} + \varrho[\lambda e_2 + (a_2 + 2te_2)\mu] g_{i0} g_{\bar{j}0} \\ & + \varrho[\lambda b_2 + (a_2 + 2tb_2)\nu] g_{i\bar{0}} g_{j0} \big\} g_{0k} \nabla y^k \wedge \nabla y^i \wedge \nabla y^j \\ & - \varrho(c_1 g_{i\bar{j}} + d_1 g_{i0} g_{\bar{j}0} + f_1 g_{i\bar{0}} g_{j0}) g_{0k} \nabla y^k \wedge dx^i \wedge dx^j . \end{aligned}$$

Thus $d\Omega_3 = \omega \wedge \Omega_3$ if and only if

$$\begin{aligned} \big\{ (2cc_2 + 2tc\sigma - c_1' + \varrho c_1) g_{i\bar{j}} g_{0k} - (d_1' - c\sigma - \varrho d_1)(g_{0i} g_{0\bar{j}} + g_{\bar{0}i} g_{0j}) g_{0k} \\ -(d_1 - cc_2)(g_{0\bar{j}} g_{ik} + g_{\bar{0}i} g_{jk} + g_{0i} g_{\bar{j}k} + g_{0j} g_{i\bar{k}}) \big\} \nabla y^k \wedge dx^i \wedge dx^j \\ +(c_2' - \sigma - \varrho c_2) g_{i\bar{j}} g_{0k} \nabla y^k \wedge \nabla y^i \wedge \nabla y^j = 0 . \end{aligned} \tag{27}$$

By (26) and (8), one finds

$$\mu = \frac{\lambda[2ca_1 - a_1'(a_1^2 + 2ct)]}{a_1(a_1^2 - 2ct)}\,, \qquad \sigma = \frac{2\lambda(c - a_1 a_1')}{a_1(a_1^2 - 2ct)}\,,$$

$$d_1 = cc_2, \quad \varrho c_2 = c_2' - \sigma, \quad \varrho d_1 = d_1' - c\sigma, \quad \varrho c_1 = c_1' - 2tc\sigma - 2cc_2.$$

Thus, (27) holds, namely,

$$d\Omega_3 = \omega \wedge \Omega_3 .$$

Moreover, from (26) and Proposition 2, it follows that J_1, J_2, J_3 are integrable. Therefore (TM, G, J_3) is a locally conformal Kähler manifold. By Proposition 3 and Proposition 4, we directly obtain that (TM, G, J_1, J_2, J_3) is a locally conformal hyper-Kähler manifold. □

Remark 4. The parameters a_1, λ are not quite arbitrary. In fact, the following conditions must be fulfilled

$$a_1 > 0\,, \qquad a_1 + 2tb_1 = \frac{a_1^2 - 2ct}{a_1 - 2ta_1'} > 0\,, \qquad a_1 + 2te_1 = \frac{a_1^2 + 2ct}{a_1} > 0\,,$$

$$\lambda > 0\,, \qquad \lambda + 2t\mu = \lambda \frac{(a_1 - 2ta_1')(a_1^2 + 2ct)}{a_1(a_1^2 - 2ct)} > 0\,.$$

Corollary 3. *Let (M, g, J) be a complex space form of holomorphic sectional curvature $4c$. If the functions b_1, e_1, λ, ν are given by*

$$b_1 = \frac{a_1 a_1' - c}{a_1 - 2ta_1'}\,, \quad e_1 = \frac{c}{a_1}\,, \quad \lambda = e^{\int \frac{2ca_1 - a_1'(a_1^2+2ct)}{a_1(a_1^2-2ct)}dt}\,, \quad \nu = 0\,, \tag{28}$$

then the almost hyper-Hermitian structure (G, J_1, J_2, J_3) is hyper-Kähler.

Proof. Under the assumption of (28), (TM, G, J_1, J_2, J_3) is a locally conformal hyper-Kähler manifold, and the Lee form $\omega = 0$. This shows that

$$d\Omega_\alpha = \omega \wedge \Omega_\alpha = 0,$$

for each α. Therefore, the almost hyper-Hermitian structure (G, J_1, J_2, J_3) on TM is hyper-Kähler. □

Example 1. Let (M, g, J) be a complex space form of holomorphic sectional curvature $4c$. If $c \geq 0$, we can consider the following functions

$$a_1 = \sqrt{e^{-2t} + 2ct}\,, \qquad b_1 = -\frac{\sqrt{e^{-2t} + 2ct}}{1+2t}\,,$$
$$e_1 = \frac{c}{\sqrt{e^{-2t} + 2ct}}\,, \qquad \nu = 0\,.$$

Clearly, all the conditions of Remark 4 and Theorem 1 are fulfilled. Therefore, we obtain a family of locally conformal hyper-Kähler structures (G, J_1, J_2, J_3) on TM. In particular, putting

$$\lambda = e^{\int \frac{e^{-2t}+4ct+c}{e^{-2t}+2ct}dt}, \tag{29}$$

we can obtain a family of hyper-Kähler structures (G, J_1, J_2, J_3) on TM. If $c < 0$, we can consider the following functions

$$a_1 = \sqrt{e^{-2t} - 2ct}\,, \quad b_1 = -\frac{(2c + e^{-2t})\sqrt{e^{-2t} - 2ct}}{(1+2t)e^{-2t}}\,,$$
$$e_1 = \frac{c}{\sqrt{e^{-2t} - 2ct}}\,, \quad \nu = 0\,.$$

Clearly, all the conditions of Remark 4 and Theorem 1 are fulfilled. Therefore, we obtain a family of locally conformal hyper-Kähler structures (G, J_1, J_2, J_3) on TM. In particular, putting

$$\lambda = e^{\int \frac{e^{-4t}+3ce^{-2t}-4c^2t}{(e^{-2t}-2ct)(e^{-2t}-4ct)}dt}, \tag{30}$$

we can obtain a family of hyper-Kähler structures (G, J_1, J_2, J_3) on TM.

Example 2. Let (M, g, J) be a complex space form of holomorphic sectional curvature $4c$. If $c \geq 0$, we can consider the following functions

$$a_1 = A + \sqrt{A^2 + 2ct}\,, \quad b_1 = e_1 = \frac{c}{A + \sqrt{A^2+2ct}}\,, \quad \lambda = B\,, \quad \mu = \nu = 0\,,$$

where A, B are two positive constants. If $c < 0$, we can consider the following functions

$$a_1 = A + \sqrt{A^2 - 2ct}\,, \quad b_1 = \frac{-c(A + 2\sqrt{A^2 - 2ct})}{A(A + \sqrt{A^2 - 2ct})}\,, \quad e_1 = \frac{c}{A + \sqrt{A^2 - 2ct}}\,,$$

$$\lambda = \frac{B}{\sqrt{A^2 - 2ct}}\,, \qquad \mu = \nu = 0\,.$$

Clearly, all the conditions of Remark 4 and Corollary 3 are fulfilled. Therefore, we obtain a family of hyper-Kähler structures (G, J_1, J_2, J_3) on TM.

In the end of the section, we compute the Levi-Civita connection $\widetilde{\nabla}$ of (TM, G). The explicit expression of $\widetilde{\nabla}$ is obtained from the following well-known formula

$$\begin{aligned} 2G(\widetilde{\nabla}_{\widetilde{X}}\widetilde{Y}, \widetilde{Z}) &= \widetilde{X}(G(\widetilde{Y}, \widetilde{Z})) + \widetilde{Y}(G(\widetilde{X}, \widetilde{Z})) - \widetilde{Z}(G(\widetilde{X}, \widetilde{Y})) \\ &\quad + G([\widetilde{X}, \widetilde{Y}], \widetilde{Z}) + G([\widetilde{Z}, \widetilde{X}], \widetilde{Y}) + G(\widetilde{X}, [\widetilde{Z}, \widetilde{Y}])\,, \end{aligned} \tag{31}$$

where $\widetilde{X}, \widetilde{Y}, \widetilde{Z} \in \mathscr{X}(TM)$. A direct computation using (5) and (31) gives the following proposition.

Proposition 5. *Let (M, g, J) be a Kähler manifold, then the Levi-Civita connection $\widetilde{\nabla}$ of (TM, G) is given by the following identities:*

$$\begin{aligned} \widetilde{\nabla}_{\partial_i}\partial_j &= Q^l_{ij}\partial_l\,, & \widetilde{\nabla}_{\delta_i}\partial_j &= \Gamma^l_{ij}\partial_l + P^l_{ji}\delta_l\,, \\ \widetilde{\nabla}_{\partial_i}\delta_j &= P^l_{ij}\delta_l\,, & \widetilde{\nabla}_{\delta_i}\delta_j &= \Gamma^l_{ij}\delta_l + S^l_{ij}\partial_l\,, \end{aligned}$$

where Q^l_{ij}, P^l_{ij}, S^l_{ij} are defined by

$$\begin{aligned} Q^l_{ij} = {} & \frac{2d_2 - c_2'}{2(c_2 + 2td_2)} g_{ij} y^l + \frac{c_2'}{2c_2}(g_{0i}\delta^l_j + g_{0j}\delta^l_i) + \frac{c_2 d_2' - 2c_2' d_2}{2c_2(c_2 + 2td_2)} g_{0i} g_{0j} y^l \\ & + \frac{c_2 f_2' - c_2' f_2 - 2f_2^2}{2c_2(c_2 + 2tf_2)}(g_{0i} g_{\bar{0}j} + g_{0j} g_{\bar{0}i}) y^{\bar{l}} + \frac{2d_2 f_2 - c_2 f_2'}{2c_2(c_2 + 2td_2)} g_{\bar{0}i} g_{\bar{0}j} y^l \\ & + \frac{f_2}{c_2}(J^l_i g_{\bar{0}j} + J^l_j g_{\bar{0}i})\,, \qquad Jy = y^k J^l_k \frac{\partial}{\partial x^l} := y^{\bar{l}} \frac{\partial}{\partial x^l}\,, \end{aligned}$$

$$\begin{aligned} P^l_{ij} = {} & \frac{c_1'}{2c_1} g_{0i}\delta^l_j + \frac{d_1}{2c_1} g_{0j}\delta^l_i + \frac{d_1}{2(c_1 + 2td_1)} g_{ij} y^l + \frac{c_1 d_1' - c_1' d_1 - d_1^2}{2c_1(c_1 + 2td_1)} g_{0i} g_{0j} y^l \\ & + \frac{c_2}{2c_1} R^l_{j0i} + \frac{c_2 d_1}{2c_1(c_1 + 2td_1)} R_{0i0j} y^l + \frac{c_1 f_1' - c_1' f_1 - f_1^2}{2c_1(c_1 + 2tf_1)} g_{0i} g_{\bar{0}j} y^{\bar{l}} \\ & + \frac{d_1 f_1}{2c_1(c_1 + 2td_1)} g_{\bar{0}i} g_{\bar{0}j} y^l + \frac{f_1}{2(c_1 + 2tf_1)} g_{\bar{i}j} y^{\bar{l}} - \frac{d_1 f_1}{2c_1(c_1 + 2tf_1)} g_{\bar{0}i} g_{0j} y^{\bar{l}} \\ & + \frac{f_1}{2c_1} g_{\bar{0}j} J^l_i + \frac{f_2}{2c_1} R^l_{j0\bar{0}} g_{\bar{0}i} + \frac{d_1 f_2}{2c_1(c_1 + 2td_1)} R_{0\bar{0}0j} g_{\bar{0}i} y^l \\ & + \frac{c_2 f_1}{2c_1(c_1 + 2tf_1)} R_{0i\bar{0}j} y^{\bar{l}} + \frac{f_1 f_2}{2c_1(c_1 + 2tf_1)} R_{0\bar{0}\bar{0}j} g_{\bar{0}i} y^{\bar{l}}\,, \end{aligned}$$

$$S_{ij}^l = -\frac{1}{2}R_{0ij}^l - \frac{c_1'}{2(c_2+2td_2)}g_{ij}y^l + \frac{2d_1d_2 - c_2d_1'}{2c_2(c_2+2td_2)}g_{0i}g_{0j}y^l$$
$$- \frac{d_1}{2c_2}(g_{0i}\delta_j^l + g_{0j}\delta_i^l) + \frac{2d_2f_1 - c_2f_1'}{2c_2(c_2+2td_2)}g_{\bar{0}i}g_{\bar{0}j}y^l$$
$$+ \frac{(d_1-f_1)f_2}{2c_2(c_2+2tf_2)}(g_{0i}g_{\bar{0}j} + g_{\bar{0}i}g_{0j})y^{\bar{l}} + \frac{f_1}{2c_2}(J_i^l g_{\bar{0}j} + g_{\bar{0}i}J_j^l)\,.$$

4 The unit tangent sphere bundle T_1M

In this section, we restrict the almost hyper-Hermitian structure $(G, J_\alpha)_{\alpha=1,2,3}$ on the unit tangent sphere bundle T_1M, obtaining an almost contact metric 3-structure $(\varphi_\alpha, \xi_\alpha, \eta_\alpha, \widehat{G})_{\alpha=1,2,3}$. We study the geometry of $(\varphi_\alpha, \xi_\alpha, \eta_\alpha, \widehat{G})_{\alpha=1,2,3}$ and find conditions under which $(\varphi_\alpha, \xi_\alpha, \eta_\alpha, \widehat{G})_{\alpha=1,2,3}$ is a Sasakian 3-structure.

Let $T_1M = \{y \in TM \colon g(y,y) = 1\}$ be the unit tangent sphere bundle on M and $\pi_1 \colon T_1M \to M$ be the canonical projection. Clearly, T_1M is a hypersurface of TM with a local expression

$$g_{ij}(x)y^iy^j = 1$$

in a local coordinates (x^i, y^i) on TM.

Let ∂_i and δ_i be as in Section 2, and define the vector field N on T_1M given by

$$N = \frac{1}{\sqrt{c_2+d_2}}y^i\partial_i\,.$$

Clearly, N is a unit normal vector field of T_1M in TM. Using this fact, one can find the tangent vector field Y_i on T_1M given by

$$Y_i = \Big(\frac{\partial}{\partial x^i}\Big)^T := \partial_i - g_{0i}y^l\partial_l\,, \quad i = 1,\dots,n\,. \tag{32}$$

Then it is easy to see that $y^iY_i = 0$. So, $Y_1,\dots,Y_n$ are not linearly independent. But we can verify that these n vectors together with $\{\delta_i\}$ span $T_y(T_1M)$ at each point $y \in T_1M$. Denote by $\widehat{G}$ the induced metric on T_1M from (TM, G). Then we have

$$\begin{cases} \widehat{G}(\delta_i,\delta_j) = c_1g_{ij} + d_1g_{0i}g_{0j} + f_1g_{\bar{0}i}g_{\bar{0}j}\,,\\ \widehat{G}(Y_i,Y_j) = c_2(g_{ij} - g_{0i}g_{0j}) + f_2g_{\bar{0}i}g_{\bar{0}j}\,,\\ \widehat{G}(\delta_i,Y_j) = 0\,,\end{cases}$$

where $c_1, c_2, d_1, d_2, f_1, f_2$ are constants because $t = \frac{1}{2}g(y,y) \equiv \frac{1}{2}$ on T_1M.

Using the almost hyper-Hermitian structure $(G, J_\alpha)_{\alpha=1,2,3}$ on TM, we can naturally construct an almost contact metric 3-structure $(\varphi_\alpha, \xi_\alpha, \eta_\alpha, \widehat{G})_{\alpha=1,2,3}$ on T_1M as follows.

First we put

$$\xi_\alpha = -J_\alpha N\,.$$

Then ξ_α is a unit tangent vector field on T_1M. Next we define a one-form η_α and a $(1,1)$ tensor field φ_α on T_1M by

$$\varphi_\alpha(V) = \tan(J_\alpha V)\,, \quad \eta_\alpha(V)N = \mathrm{nor}(J_\alpha V)\,, \quad \forall\, V \in TT_1M\,,$$

where tan: $TTM \to TT_1M$, nor: $TTM \to T^\perp T_1M$ are the usual projection operators. Here $T^\perp T_1M$ denotes the normal bundle of T_1M in TM.

By the above identities, one easily verifies that the following relations hold.

$$\varphi_\alpha^2 = -\mathrm{id} + \eta_\alpha \otimes \xi_\alpha\,, \quad \varphi_\alpha(\xi_\alpha) = 0\,, \quad \eta_\alpha \circ \varphi_\alpha = 0\,, \quad \eta_\alpha(\xi_\alpha) = 1\,,$$
$$\eta_\alpha(V) = \widehat{G}(V, \xi_\alpha)\,, \quad \widehat{G}(\varphi_\alpha V, \varphi_\alpha W) = \widehat{G}(V,W) - \eta_\alpha(V)\eta_\alpha(W)\,,$$

for all vector fields V, W on T_1M. This shows that $(\varphi_\alpha, \xi_\alpha, \eta_\alpha, \widehat{G})$ is an almost contact metric structure. Moreover,

$$\begin{gathered}\varphi_\gamma = \varphi_\alpha\varphi_\beta - \eta_\beta \otimes \xi_\alpha = -\varphi_\beta\varphi_\alpha + \eta_\alpha \otimes \xi_\beta\,, \qquad \eta_\alpha(\xi_\beta) = \delta_{\alpha\beta}\,, \\ \xi_\gamma = \varphi_\alpha\xi_\beta = -\varphi_\beta\xi_\alpha\,, \qquad \eta_\gamma = \eta_\alpha \circ \varphi_\beta = -\eta_\beta \circ \varphi_\alpha\,,\end{gathered} \tag{33}$$

where (α, β, γ) is a cyclic permutation of $(1, 2, 3)$. Thus we have proved the following result.

Proposition 6. *$(\varphi_\alpha, \xi_\alpha, \eta_\alpha, \widehat{G})_{\alpha=1,2,3}$ is an almost contact metric 3-structure on T_1M.*

The almost contact metric 3-structure $(T_1M, \varphi_\alpha, \xi_\alpha, \eta_\alpha, \widehat{G})_{\alpha=1,2,3}$ can be locally expressed as

$$\xi_1 = \frac{1}{\sqrt{c_1+d_1}} y^k \delta_k\,, \qquad \eta_1(\delta_i) = \sqrt{c_1+d_1}\; g_{0i}\,, \qquad \eta_1(Y_i) = 0\,,$$
$$\varphi_1(\delta_i) = (a_1\delta_i^k + e_1 g_{\bar{0}i} y^{\bar{k}})Y_k\,, \qquad \varphi_1(Y_i) = -[a_2(\delta_i^k - g_{0i}y^k) + e_2 g_{\bar{0}i} y^{\bar{k}}]\delta_k\,,$$
$$\xi_2 = -\frac{1}{\sqrt{c_1+d_1}} y^{\bar{k}} \delta_k\,, \qquad \eta_2(\delta_i) = \sqrt{c_1+d_1}\; g_{0\bar{i}}\,, \qquad \eta_2(Y_i) = 0\,,$$
$$\varphi_2(\delta_i) = (a_1 J_i^k + e_1 g_{0i} y^{\bar{k}})Y_k\,, \qquad \varphi_2(Y_i) = [a_2(J_i^k - g_{0i}y^{\bar{k}}) + e_2 g_{0\bar{i}} y^k]\delta_k\,,$$
$$\xi_3 = -\frac{1}{\sqrt{c_2+f_2}} y^{\bar{k}} Y_k\,, \qquad \eta_3(\delta_i) = 0, \qquad \eta_3(Y_i) = \sqrt{c_2+f_2}\; g_{0\bar{i}}\,,$$
$$\varphi_3(\delta_i) = -J_i^k\delta_k\,, \qquad \varphi_3(Y_i) = (J_i^k - g_{0i}y^{\bar{k}})Y_k\,.$$

Furthermore, we find from (32) and (5) that

$$[Y_i, Y_j] = -g_{0j}Y_i + g_{0i}Y_j\,, \qquad [\delta_i, Y_j] = \Gamma_{ij}^k Y_k\,, \qquad [\delta_i, \delta_j] = -R_{0ij}^k Y_k\,.$$

Then we have

$$d\eta_1(\delta_i, \delta_j) = 0\,, \quad d\eta_1(Y_i, Y_j) = 0\,, \quad d\eta_1(\delta_i, Y_j) = -\sqrt{c_1+d_1}(g_{ij} - g_{0i}g_{0j})\,,$$
$$\widehat{G}(\delta_i, \varphi_1 Y_j) = -\lambda(g_{ij} - g_{0i}g_{0j}) - \nu g_{\bar{0}i}g_{\bar{0}j}\,, \quad \widehat{G}(\delta_i, \varphi_1\delta_j) = \widehat{G}(Y_i, \varphi_1 Y_j) = 0\,,$$
$$d\eta_2(\delta_i, \delta_j) = -\sqrt{c_1+d_1}\left\{F\Big(\frac{\partial}{\partial x^i}, y, \frac{\partial}{\partial x^j}\Big) - F\Big(\frac{\partial}{\partial x^j}, y, \frac{\partial}{\partial x^i}\Big)\right\},$$
$$d\eta_2(\delta_i, Y_j) = \sqrt{c_1+d_1}(g_{i\bar{j}} - g_{0j}g_{\bar{0}i})\,, \quad d\eta_2(Y_i, Y_j) = 0\,,$$

$$\widehat{G}(\delta_i,\varphi_2 Y_j)=\lambda(g_{i\bar{j}}-g_{\bar{0}i}g_{0j})+\nu g_{0i}g_{0\bar{j}}\,,\quad \widehat{G}(\delta_i,\varphi_2\delta_j)=\widehat{G}(Y_i,\varphi_2 Y_j)=0\,,$$

$$d\eta_3(\delta_i,\delta_j)=\sqrt{c_2+f_2}R^k_{0ij}g_{\bar{k}0}\,,\quad d\eta_3(\delta_i,Y_j)=\sqrt{c_2+f_2}F\Big(\frac{\partial}{\partial x^i},\frac{\partial}{\partial x^j},y\Big)\,,$$

$$d\eta_3(Y_i,Y_j)=2\sqrt{c_2+f_2}(g_{i\bar{j}}-g_{i0}g_{\bar{j}0}+g_{j0}g_{\bar{i}0})\,,\quad \widehat{G}(\delta_i,\varphi_3 Y_j)=0\,,$$

$$\widehat{G}(\delta_i,\varphi_3\delta_j)=c_1 g_{\bar{i}j}+d_1(g_{i0}g_{\bar{0}j}-g_{i\bar{0}}g_{j0})\,,$$

$$\widehat{G}(Y_i,\varphi_3 Y_j)=c_2(g_{i\bar{j}}-g_{i0}g_{\bar{j}0}-g_{0j}g_{i\bar{0}})\,,$$

where $F(\frac{\partial}{\partial x^i},\frac{\partial}{\partial x^j},y)=g\big((\nabla_{\frac{\partial}{\partial x^i}}J)\frac{\partial}{\partial x^j},y\big)$.

In order to get a contact metric 3-structure on T_1M, i.e.

$$\widehat{G}(V,\varphi_\alpha W)=\frac{1}{2}d\eta_\alpha(V,W),\quad \forall\, V,W\in\mathscr{X}(T_1M),$$

we have to modify the almost contact metric 3-structure in the following way (see e.g. [2]):

$$\varphi_\alpha^{\rm new}=\varphi_\alpha\,,\quad \xi_\alpha^{\rm new}=\frac{2\lambda}{\sqrt{c_1+d_1}}\xi_\alpha\,,\quad \eta_\alpha^{\rm new}=\frac{\sqrt{c_1+d_1}}{2\lambda}\eta_\alpha\,,\quad \widehat{G}^{\rm new}=\frac{c_1+d_1}{4\lambda^2}\widehat{G}\,.$$

In the next discussion, we still denote by $(\varphi_\alpha,\xi_\alpha,\eta_\alpha,\widehat{G})$ the new almost contact metric 3-structure, then we obtain immediately

Proposition 7. *$(\varphi_\alpha,\xi_\alpha,\eta_\alpha,\widehat{G})_{\alpha=1,2,3}$ is a contact metric 3-structure on T_1M if and only if (M,g,J) is a Kähler manifold and*

$$\nu=0\,,\qquad a_1=e_1\,,\qquad R_{0\bar{0}ij}=2a_1^2(g_{i\bar{j}}+g_{i\bar{0}}g_{j0}-g_{j\bar{0}}g_{i0})\,. \tag{34}$$

Definition 3. A contact metric 3-structure $(\varphi_\alpha,\xi_\alpha,\eta_\alpha,\widehat{G})_{\alpha=1,2,3}$ is called a Sasakian 3-structure if each ξ_α is a Killing vector field and

$$(\widehat{\nabla}_V\varphi_\alpha)W=\widehat{G}(V,W)\xi_\alpha-\eta_\alpha(W)V\,,$$

for all vector fields V,W.

By Proposition 5, we can obtain the following proposition.

Proposition 8. *Let (M,g,J) be a Kähler manifold and denote by $\widehat{\nabla}$ the Levi-Civita connection of $(T_1M,\widehat{G})$, then*

$$\widehat{\nabla}_{\delta_i}\delta_j=\Gamma^k_{ij}\delta_k-\left(\frac{d_1}{2c_2}\left(\delta^k_i g_{0j}+\delta^k_j g_{0i}\right)+\frac{1}{2}R^k_{0ij}+\frac{f_1}{2c_2}(g_{0\bar{j}}J^k_i+g_{0\bar{i}}J^k_j)\right)Y_k\,,$$

$$
\begin{aligned}
\widehat{\nabla}_{\delta_i} Y_j = \Gamma_{ij}^k Y_k + \frac{1}{2}\bigg(& \frac{d_1}{c_1+d_1} g_{ij} y^k + \frac{d_1}{c_1} g_{0i}\delta_j^k + \frac{c_2}{c_1} R_{i0j}^k - \frac{d_1(2c_1+d_1)}{c_1(c_1+d_1)} g_{0i} g_{0j} y^k \\
& - \frac{c_2 d_1}{c_1(c_1+d_1)} R_{0ji0} y^k + \frac{d_1 f_1}{c_1(c_1+d_1)} g_{\bar{0}i} g_{\bar{0}j} y^k + \frac{f_1}{c_1+f_1} g_{i\bar{j}} y^{\bar{k}} \\
& - \frac{d_1 f_1}{c_1(c_1+f_1)} g_{0i} g_{\bar{0}j} y^{\bar{k}} + \frac{f_1}{c_1} g_{\bar{0}i} J_j^k + \frac{f_2}{c_1} R_{i0\bar{0}}^k g_{\bar{0}j} \\
& + \frac{d_1 f_2}{c_1(c_1+d_1)} R_{0\bar{0}0i} g_{\bar{0}j} y^k - \frac{f_1(2c_1+f_1)}{c_1(c_1+f_1)} g_{\bar{0}i} g_{0j} y^{\bar{k}} \\
& + \frac{c_2 f_1}{c_1(c_1+f_1)} R_{0j\bar{0}i} y^{\bar{k}} + \frac{f_1 f_2}{c_1(c_1+f_1)} R_{0\bar{0}\bar{0}i} g_{\bar{0}j} y^{\bar{k}} \bigg)\delta_k\,, \\
\widehat{\nabla}_{Y_i} Y_j = -g_{0j} Y_i - & \frac{f_2}{c_2}\left(g_{0\bar{j}} J_i^k + g_{0\bar{i}} J_j^k - g_{0\bar{j}} g_{0i} y^{\bar{k}} - g_{0j} g_{0\bar{i}} y^{\bar{k}} \right) Y_k\,, \\
\widehat{\nabla}_{Y_i} \delta_j = \frac{1}{2}\bigg(& \frac{d_1}{c_1+d_1} g_{ij} y^k - \frac{d_1(2c_1+d_1)}{c_1(c_1+d_1)} g_{0i} g_{0j} y^k - \frac{c_2 d_1}{c_1(c_1+d_1)} R_{0ij0} y^k \\
& + \frac{c_2}{c_1} R_{j0i}^k + \frac{d_1}{c_1} g_{0j}\delta_i^k + \frac{d_1 f_1}{c_1(c_1+d_1)} g_{\bar{0}i} g_{\bar{0}j} y^k + \frac{f_1}{c_1+f_1} g_{\bar{i}j} y^{\bar{k}} \\
& - \frac{d_1 f_1}{c_1(c_1+d_1)} g_{\bar{0}i} g_{0j} y^{\bar{k}} + \frac{f_1}{c_1} g_{\bar{0}j} J_i^k + \frac{f_2}{c_1} R_{j0\bar{0}}^k g_{\bar{0}i} + \frac{d_1 f_2}{c_1(c_1+d_1)} R_{0\bar{0}0j} g_{\bar{0}i} y^k \\
& - \frac{f_1(2c_1+f_1)}{c_1(c_1+f_1)} g_{0i} g_{\bar{0}j} y^{\bar{k}} + \frac{c_2 f_1}{c_1(c_1+f_1)} R_{0i\bar{0}j} y^{\bar{k}} + \frac{f_1 f_2}{c_1(c_1+f_1)} R_{0\bar{0}\bar{0}j} g_{\bar{0}i} y^{\bar{k}} \bigg)\,\delta_k.
\end{aligned}
$$

Remark 5. If the base manifold (M, g, J) is a complex space form of holomorphic sectional curvature $4c > 0$ and the parameters a_1, e_1, ν satisfy

$$
a_1^2 = c\,, \qquad a_1 = e_1\,, \qquad \nu = 0\,, \tag{35}
$$

then (34) holds. In this case, $(\varphi_\alpha, \xi_\alpha, \eta_\alpha, \widehat{G})_{\alpha=1,2,3}$ is a contact metric 3-structure on T_1M. From (35) and (8), it follows that

$$
c_1 = d_1 = f_1 = \lambda a_1\,, \qquad f_2 = \lambda e_2\,, \qquad e_2 = -\frac{1}{2a_1}\,.
$$

By Proposition 8 and Remark 5, we have the following

Proposition 9. *Let (M, g, J) be a complex space form of positive holomorphic sectional curvature $4c$. If the parameters a_1, e_1, ν satisfy (35), then the Levi-Civita connection of the metric $\widehat{G}$ is determined by the following equations.*

$$
\begin{aligned}
\widehat{\nabla}_{\delta_i} \delta_j &= \Gamma_{ij}^k \delta_k + c\left(-g_{0j}\delta_i^k + g_{\bar{0}j} J_i^k - g_{i\bar{j}} y^{\bar{k}} \right) Y_k \\
\widehat{\nabla}_{\delta_i} Y_j &= \frac{1}{2}\left(g_{ij} y^k - g_{0i} g_{0j} y^k + g_{i\bar{j}} y^{\bar{k}} - g_{\bar{0}j} J_i^k + g_{0i} g_{\bar{0}j} y^{\bar{k}} - g_{\bar{0}i} g_{\bar{0}j} y^k - g_{\bar{0}i} g_{0j} y^{\bar{k}} \right) \delta_k \\
&\quad + \Gamma_{ij}^k Y_k \\
\widehat{\nabla}_{Y_i} Y_j &= -g_{0j} Y_i - \frac{1}{2}\left(g_{\bar{0}j} J_i^k + g_{\bar{0}i} J_j^k - g_{\bar{0}i} g_{0j} y^{\bar{k}} - g_{0i} g_{\bar{0}j} y^{\bar{k}} \right) Y_k \\
\widehat{\nabla}_{Y_i} \delta_j &= \frac{1}{2}\left(g_{ij} y^k - g_{0i} g_{0j} y^k + g_{\bar{i}j} y^{\bar{k}} - g_{\bar{0}i} J_j^k + g_{\bar{0}i} g_{0j} y^{\bar{k}} - g_{\bar{0}i} g_{\bar{0}j} y^k - g_{0i} g_{\bar{0}j} y^{\bar{k}} \right) \delta_k
\end{aligned}
$$

Remark 6. Under the assumption of Proposition 9, a direct calculation shows that the following relations hold:

$$\widehat{G}(\widehat{\nabla}_V \xi_\alpha, W) = -\widehat{G}(\widehat{\nabla}_W \xi_\alpha, V)\,,$$
$$(\widehat{\nabla}_V \varphi_\alpha) W = \widehat{G}(V,W)\xi_\alpha - \eta_\alpha(W) V\,,$$

for each α and all vector fields V, W on T_1M. Therefore, we have proved the following

Theorem 2. *There exists a class of Sasakian 3-structures on the unit tangent sphere bundle of a complex space form of positive holomorphic sectional curvature.*

Acknowledgement

The authors would like to express their gratitude to Professor Qing-Ming Cheng for many valuable discussions. Besides, the authors want to thank the referee for his/her helpful comments on the paper.

This work is supported by grants Nos. U1304101 and 11171091 of NSFC and NSF of Henan Povince (No. 132300410141).

References

[1] M. Anastasiei: Locally conformal Kähler structures on tangent manifold of a space form. *Libertas Math.* 19 (1999) 71–76.

[2] D.E. Blair: *Riemannian geometry of contact and symplectic manifolds.* Progr. Math. Birkhäuser, Boston (2002).

[3] S.A. Bogdanovich, A.A. Ermolitski: On almost hyperHermitian structures on Riemannian manifolds and tangent bundles. *Cent. Eur. J. Math.* 2 (5) (2004) 615–623.

[4] E. Calabi: Métriques kähleriennes et fibrés holomophes. *Ann. Sci. École Norm. Sup.* 12 (1979) 269–294.

[5] J. Cheeger, D. Gromoll: On the structure of complete manifolds of nonnegative curvature. *Ann. Math.* 96 (1972) 413–443.

[6] P. Dombrowski: On the geometry of the tangent bundle. *J. Reine Angew. Math.* 210 (1962) 73–88.

[7] O. Kowalski: Curvature of the induced Riemannian metric on the tangent bundle of a Riemannian manifold. *J. Reine Angew. Math.* 250 (1971) 124–129.

[8] X.X. Li, X.R. Qi: A note on some metrics on tangent bundles and unit tangent sphere bundles. *J. Math. Res. Exposition* 28 (4) (2008) 829–838.

[9] M.I. Munteanu: Some aspects on the geometry of the tangent bundles and tangent sphere bundles of a Riemannian manifold. *Mediterr. J. Math.* 5 (2008) 43–59.

[10] E. Musso, F. Tricerri: Riemannian metrics on tangent bundles. *Ann. Mat. Pura Appl.* 150 (4) (1988) 1–19.

[11] T. Nagano: Isometries on complex-product spaces. *Tensor* 9 (1959) 47–61.

[12] V. Oproiu: A Kähler Einstein structure on the tangent bundle of a space form. *Int. J. Math. Math. Sci.* 25 (2001) 183–195.

[13] V. Oproiu: Hyper-Kähler structures on the tangent bundle of a Kähler manifold. *Balkan J. Geom. Appl.* 15 (1) (2010) 104–119.

[14] V. Oproiu, N. Papaghiuc: General natural Einstein Kähler structures on tangent bundles. *Differential Geom. Appl.* 27 (2009) 384–392.

[15] V. Oproiu, D.D. Poroşniuc: A class of Kähler Einstein structures on the cotangent bundle. *Publ. Math. Debrecen* 66 (3–4) (2005) 457–478.

[16] L. Ornea, P. Piccinni: Locally conformal Kähler structures in quaternionic geometry. *Trans. Amer. Math. Soc.* 349 (2) (1997) 641–655.

[17] D.D. Poroşniuc: A class of locally symmetric Kähler Einstein structures on the nonzero cotangent bundle of a space form. *Balkan J. Geom. Appl.* 9 (2) (2004) 68–81.

[18] S. Sasaki: On the differential geometry of tangent bundles of Riemannian manifolds. *Tôhoku Math. J.* 10 (1958) 338–354.

[19] S. Tachibana, M. Okumura: On the almost-complex structure of tangent bundles of Riemannian spaces. *Tôhoku Math. J.* 14 (2) (1962) 156–161.

[20] M. Tahara, L. Vanhecke, Y. Watanabe: New structures on tangent bundles. *Note Mat.* 18 (1) (1998) 131–141.

[21] I. Vaisman: On locally conformal almost Kähler manifolds. *Israel J. Math.* 24 (1976) 338–351.

[22] B.V. Zayatuev: On a class of almost-Hermitian structures on tangent bundles. *Math. Notes* 76 (5) (2004) 682–688.

Authors' addresses:
Xuerong Qi: Department of Mathematics, Zhengzhou University, Zhengzhou 450001, P. R. China

E-mail: xrqi@zzu.edu.cn

Linfen Cao, Xingxiao Li: Department of Mathematics, Henan Normal University, Xinxiang 453007, P. R. China

E-mail: caolf2010@yahoo.com, xxl@henannu.edu.cn

Received: 7 November, 2013
Accepted for publication: 31 January, 2014
Communicated by: Haizhong Li

Communications in Mathematics 22 (2014) 31–48

Stratonovich-Weyl correspondence for the Jacobi group

Benjamin Cahen

Abstract. We construct and study a Stratonovich-Weyl correspondence for the holomorphic representations of the Jacobi group.

1 Introduction

The notion of Stratonovich-Weyl correspondence was introduced in [27] as a generalization of the classical Weyl correspondence [1]. The systematic study of the Stratonovich-Weyl correspondences began with the work of J.M. Gracia-Bondìa, J.C. Vàrilly and their co-workers (see [15], [17], [19] and [20]).

Definition 1. [19] Let G be a Lie group and π a unitary representation of G on a Hilbert space $\mathcal{H}$. Let M be a homogeneous G-space and μ a (suitably normalized) G-invariant measure on M. Then a Stratonovich-Weyl correspondence for the triple (G, π, M) is an isomorphism W from a vector space of operators on $\mathcal{H}$ to a space of (generalized) functions on M satisfying the following properties:

1. W maps the identity operator of $\mathcal{H}$ to the constant function 1;
2. the function $W(A^*)$ is the complex-conjugate of $W(A)$;
3. Covariance: we have $W(\pi(g)\, A\, \pi(g)^{-1})(x) = W(A)(g^{-1} \cdot x)$;
4. Traciality: we have

$$\int_M W(A)(x) W(B)(x)\, d\mu(x) = \mathrm{Tr}(AB).$$

2010 MSC: 22E27, 32M10, 46E22

Key words: Berezin quantization, Berezin transform, quasi-Hermitian Lie group, unitary representation, holomorphic representation, reproducing kernel Hilbert space, Jacobi group, Stratonovich-Weyl correspondence

The basic example is the case when G is the $(2n+1)$-dimensional Heisenberg group H_n acting on $\mathbb{R}^{2n}$ by translations and π is a Schrödinger representation of H_n on $L^2(\mathbb{R}^n)$. In this case, the classical Weyl correspondence gives a Stratonovich--Weyl correspondence for the triple $(H_n, \pi, \mathbb{R}^{2n})$ [18], [19].

Stratonovich-Weyl correspondences were constructed for various Lie group representations, in particular for the massive representations of the Poincaré group [15], [19].

In [13], we constructed and studied a Stratonovich-Weyl correspondence for a quasi-Hermitian Lie group G and a unitary representation π of G which is holomorphically induced from a unitary character of a compactly embedded subgroup K of G (see also our earlier papers [11] and [12]). The construction is based on an idea of [17] consisting in modifying suitably the Berezin correspondence S (see also [2] and [3]). More precisely, recall that S is an isomorphism from the Hilbert space of all Hilbert-Schmidt operators on $\mathcal{H}$ (endowed with the Hilbert-Schmidt norm) onto a space of square integrable functions on a homogeneous complex domain [28]. Then the Stratonovich-Weyl correspondence W is obtained by taking the isometric part in the polar decomposition of S, that is, $W := (SS^*)^{-1/2}S$. Note that $B := SS^*$ is the so-called Berezin transform which have been intensively studied by many authors, see in particular [16], [23], [24], [28], [29].

In [13], we showed that, when the Lie algebra $\mathfrak{g}$ of G is reductive, the mappings B and W can be extended to a class of functions which contains $S(d\pi(X))$ for each $X \in \mathfrak{g}$ and that, for each simple ideal $\mathfrak{s}$ in $\mathfrak{g}$, there exists a constant $c \geq 0$ such that $W(d\pi(X)) = cS(d\pi(X))$ for each $X \in \mathfrak{s}$. However, it seems difficult to obtain the analogous results in the general case.

In the present paper, we aim to study B and W in the particular case of the Jacobi group. The Jacobi group is the semi-direct product of the 3-dimensional real Heisenberg group by the unitary group $SU(1,1)$. This group plays an important role in different areas of Mathematics and Physics as for instance Theory of automorphic forms (Jacobi forms, theta functions), Quantum Optics (squeezed states) and Harmonic Analysis, see [4] and [9]. In particular, the Jacobi group appears as an important example of non-reductive Lie group of Harish-Chandra type [22], [26] and its holomorphic unitary representations were studied by many authors, see [4], [5], [8], [9] and [22]. Moreover, the metaplectic factorization [22] should be used to reduce, in some sense, the general case to the case of some generalized Jacobi group and then the study of the case of the Jacobi group can be considered as a first step towards the general case. Recall that the metaplectic factorization is a method for decomposing a unitary highest weight representation of a quasi-Hermitian Lie group as the tensor product of a unitary highest weight representation of a reductive group and a highest weight representation of some generalized Heisenberg group [22], p. 361.

In this paper, we begin by some generalities on the Jacobi group (Section 2) and its holomorphic representations (Section 3). We introduce the Berezin correspondence S, the Berezin transform B and the Stratonovich-Weyl correspondence W (Section 4). Under some technical assumptions, we extend B to a class of functions which contains $S(d\pi(X))$ for each $X \in \mathfrak{g}$ (Section 5). Finally, we give an explicit expression for $W(d\pi(X))$, $X \in \mathfrak{g}$ (Section 6).

2 The Jacobi group

This section is devoted to generalities on the Jacobi group. The material of this section and of the following section is essentially taken from [18], Chapter 4, [22], Chapters VII and XII, [14] (see also [9]).

Consider the symplectic form ω on $\mathbb{C}^2$ defined by

$$\omega\big((z,w),(z',w')\big)=\frac{i}{2}(zw'-z'w).$$

for $z,w,z',w'\in\mathbb{C}$. The 3-dimensional real Heisenberg group is

$$H:=\big\{\big((z,\bar z),c\big):z\in\mathbb{C},\,c\in\mathbb{R}\big\}$$

endowed with the multiplication

$$\big((z,\bar z),c\big)\cdot\big((z',\bar z'),c'\big)=\Big((z+z',\bar z+\bar z'),c+c'+\tfrac{1}{2}\omega\big((z,\bar z),(z',\bar z')\big)\Big).$$

Then the complexification H^c of H is $H^c:=\{((z,w),c):z,w\in\mathbb{C},c\in\mathbb{C}\}$ and the multiplication of H^c is obtained by replacing $(z,\bar z)$ by (z,w) and $(z',\bar z')$ by (z',w') in the previous formula. We denote by $\mathfrak{h}$ and $\mathfrak{h}^c$ the Lie algebras of H and H^c.

Now consider the group $S:=SU(1,1)$ consisting of all matrices

$$h=\begin{pmatrix} a & b\\ \bar b & \bar a\end{pmatrix},\quad a,b\in\mathbb{C}\,,\quad |a|^2-|b|^2=1\,.$$

The group S acts on H by

$$h\cdot\big((z,\bar z),c\big)=\big(h(z,\bar z),c\big)=(az+b\bar z,\bar a\bar z+\bar b z,c)$$

for h as before, the elements of $\mathbb{C}^2$ being considered as column vectors. Then we can form the semi-direct product $G:=H\rtimes S$ which is called the Jacobi group. The elements of G can be written as $((z,\bar z),c,h)$ where $z\in\mathbb{C}$, $c\in\mathbb{R}$ and $h\in S$. The multiplication of G is thus given by

$$\big((z,\bar z),c,h\big)\cdot\big((z',\bar z'),c',h'\big)=\Big((z,\bar z)+h(z',\bar z'),c+c'+\tfrac{1}{2}\omega\big((z,\bar z),h(z',\bar z')\big),hh'\Big).$$

The complexification of S is $S^c=SL(2,\mathbb{C})$. The complexification G^c of G is then the semi-direct product $G^c=H^c\rtimes SL(2,\mathbb{C})$ and the multiplication of G^c is obtained by replacing $\bar z$ and $\bar z'$ by w and w' in the preceding formula. We denote by $\mathfrak{s}$, $\mathfrak{s}^c$, $\mathfrak{g}$ and $\mathfrak{g}^c$ the Lie algebras of S, S^c, G and G^c. The Lie bracket of $\mathfrak{g}^c$ is given by

$$\Big[\big((z,w),c,A\big),\big((z',w'),c',A'\big)\Big]=\big(A(z',w')-A'(z,w),\omega((z,w),(z',w')),[A,A']\big).$$

Let θ denote conjugation with respect to the real form $\mathfrak{g}$ of $\mathfrak{g}^c$. For $X\in\mathfrak{g}^c$, we set $X^*=-\theta(X)$. We can easily verify that if $X=\big((z,w),c,\left(\begin{smallmatrix} a & b\\ c & -a\end{smallmatrix}\right)\big)\in\mathfrak{g}^c$ then

$$X^*=\left((-\bar w,-\bar z),-\bar c,\begin{pmatrix}\bar a & -\bar c\\ -\bar b & -\bar a\end{pmatrix}\right).$$

Also, we denote by $g \to g^*$ the involutive anti-automorphism of G^c which is obtained by exponentiating $X \to X^*$ to G^c.

Let K be the subgroup of G consisting of all elements $\big((0,0), c, \left(\begin{smallmatrix} a & 0 \\ 0 & \bar{a} \end{smallmatrix}\right)\big)$ where $c \in \mathbb{R}$ and $|a| = 1$. Then the Lie algebra $\mathfrak{k}$ of K is a maximal compactly embedded (Cartan) subalgebra of $\mathfrak{g}$. Let us introduce the linear form ε defined on $\mathfrak{k}$ by $\varepsilon\big((0,0), c, \left(\begin{smallmatrix} a & 0 \\ 0 & \bar{a} \end{smallmatrix}\right)\big) = a$. Then one can easily verify that the roots of $\mathfrak{k}^c$ on $\mathfrak{g}^c$ are $\pm\varepsilon$ and $\pm 2\varepsilon$. In the terminology of [22] p. 234–235, $\pm 2\varepsilon$ are the (non-compact) semi-simple roots, $\pm\varepsilon$ are the solvable roots (in that case, there is no compact roots). As in [22] p. 249, we can choose the adapted system of positive roots to be $\{\varepsilon, 2\varepsilon\}$. The root space decomposition of $\mathfrak{g}^c$ is then $\mathfrak{g}^c = \mathfrak{p}^+ \oplus \mathfrak{k}^c \oplus \mathfrak{p}^-$ where

$$\mathfrak{p}^+ = \mathfrak{g}_\varepsilon \oplus \mathfrak{g}_{2\varepsilon} = \left\{ \left((z,0), 0, \begin{pmatrix} 0 & u \\ 0 & 0 \end{pmatrix} \right) \,:\, z \in \mathbb{C}, u \in \mathbb{C} \right\}$$

and

$$\mathfrak{p}^- = \mathfrak{g}_{-\varepsilon} \oplus \mathfrak{g}_{-2\varepsilon} = \left\{ \left((0,w), 0, \begin{pmatrix} 0 & 0 \\ v & 0 \end{pmatrix} \right) \,:\, w \in \mathbb{C}, v \in \mathbb{C} \right\}.$$

In the rest of the paper, we denote by $a(z,u)$ the element $\big((z,0), 0, \left(\begin{smallmatrix} 0 & u \\ 0 & 0 \end{smallmatrix}\right)\big)$ of $\mathfrak{p}^+$. Also, we denote by $p_{\mathfrak{p}^+}$, $p_{\mathfrak{k}^c}$ and $p_{\mathfrak{p}^-}$ the projections of $\mathfrak{g}^c$ onto $\mathfrak{p}^+$, $\mathfrak{k}^c$ and $\mathfrak{p}^-$ associated with the above direct decomposition.

Let P^+ and P^- be the analytic subgroups of G^c with Lie algebras $\mathfrak{p}^+$ and $\mathfrak{p}^-$. Then we have

$$P^+ = \left\{ \left((z,0), 0, \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} \right) \,:\, z \in \mathbb{C}, u \in \mathbb{C} \right\}$$

and

$$P^- = \left\{ \left((0,w), 0, \begin{pmatrix} 1 & 0 \\ v & 1 \end{pmatrix} \right) \,:\, w \in \mathbb{C}, v \in \mathbb{C} \right\}.$$

Note that G is a group of the Harish-Chandra type [22], p. 507, that is, the following properties are satisfied:

1. $\mathfrak{g}^c = \mathfrak{p}^+ \oplus \mathfrak{k}^c \oplus \mathfrak{p}^-$ is a direct sum of vector spaces, $(\mathfrak{p}^+)^* = \mathfrak{p}^-$ and $[\mathfrak{k}^+, \mathfrak{p}^\pm] \subset \mathfrak{p}^\pm$;

2. The multiplication map $P^+K^cP^- \to G^c$, $(z,k,y) \to zky$ is a biholomorphic diffeomorphism onto its open image;

3. $G \subset P^+K^cP^-$ and $G \cap K^cP^- = K$.

We denote by $\zeta : P^+K^cP^- \to P^+$, $\kappa : P^+K^cP^- \to K^c$ and $\eta : P^+K^cP^- \to P^-$ the projections onto P^+-, K^c- and P^--components. We can verify that $g = \big((z_0, w_0), c_0, \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\big)$ in G^c has a $P^+K^cP^-$-decomposition

$$g = \left((z,0), 0, \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} \right) \cdot \left((0,0), \tilde{c}, \begin{pmatrix} p & 0 \\ 0 & 1/p \end{pmatrix} \right) \cdot \left((0,w), 0, \begin{pmatrix} 1 & 0 \\ v & 1 \end{pmatrix} \right)$$

if and only if $d \neq 0$ and, in this case, we have $z = z_0 - bd^{-1}w_0$, $u = bd^{-1}$, $v = cd^{-1}$, $w = d^{-1}w_0$, $p = d^{-1}$ and $\tilde{c} = c_0 - (1/4)i(z_0 - bd^{-1}w_0)w_0$.

Now we introduce an action (defined almost everywhere) of G^c on $\mathfrak{p}^+$. For $Z \in \mathfrak{p}^+$ and $g \in G^c$ with $g \exp Z \in P^+K^cP^-$, we define the element $g \cdot Z$ of $\mathfrak{p}^+$ by $g \cdot Z := \log \zeta(g \exp Z)$. From the above formula for the $P^+K^cP^-$-decomposition, we deduce that the action of $g = \big((z_0, w_0), c_0, \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\big) \in G^c$ on $a(z,u) \in \mathfrak{p}^+$ is given by $g \cdot a(z,u) = a(z', u')$ where

$$u' = (au+b)(cu+d)^{-1}$$

and

$$z' = z_0 + az - (au+b)(cu+d)^{-1}(w_0 + cz)\,.$$

This implies that

$$\mathcal{D} := G \cdot 0 = \{a(z,u) \in \mathfrak{p}^+ \;:\; |u| < 1\} \cong \mathbb{C} \times \mathbb{D}$$

where $\mathbb{D}$ denotes the unit open disk of $\mathbb{C}$.

We can easily obtain a useful section $Z \to g_Z$ for the action of G on $\mathcal{D}$. Let $Z = a(z,u) \in \mathcal{D}$. Define $g_Z := ((z_0, \bar{z}_0), 0, h_0) \in G$ as follows. We set

$$z_0 = (1 - u\bar{u})^{-1}(z + u\bar{z})\,, \quad a = (1-u\bar{u})^{-1/2}, \quad b = (1-u\bar{u})^{-1/2}u$$

and $h_0 = \left(\begin{smallmatrix} a & b \\ \bar{b} & \bar{a} \end{smallmatrix}\right)$. Then one has $g_Z \cdot 0 = Z$.

From this, we can compute the G-invariant measure μ on $\mathcal{D}$. Let $d\mu_L(z,u)$ be the Lebesgue measure on $\mathbb{C} \times \mathbb{D}$ normalized as follows. If $z = x + iy$ and $u = v + iw$ with $x, y, v, w \in \mathbb{R}$ then $d\mu_L(z,u) := dx\,dy\,dv\,dw$. Thus, we easily obtain that $d\mu(Z) = (1 - u\bar{u})^{-3}\, d\mu_L(z,u)$. This result can be also deduced from the general formula for the invariant measure, see [22], p. 538.

Now we compute the adjoint and coadjoint actions of G^c. This can be done as follows. We begin by the adjoint action of G^c. Let $g = (v_0, c_0, h_0) \in G^c$ where $v_0 \in \mathbb{C}^2$, $c_0 \in \mathbb{C}$ and $h_0 \in S^c = SL(2,\mathbb{C})$ and $X = (w, c, U) \in \mathfrak{g}^c$ where $w \in \mathbb{C}^2$, $c \in \mathbb{C}$ and $U \in \mathfrak{s}^c$. We set $\exp(tX) = (w(t), c(t), \exp(tU))$. Then, since the derivatives of $w(t)$ and $c(t)$ at $t = 0$ are w and c, we find that

$$\begin{aligned} \mathrm{Ad}(g)X &= \frac{d}{dt}(g \exp(tX) g^{-1})|_{t=0} \\ &= \big(h_0 w - (\mathrm{Ad}(h_0)U)v_0, c + \omega(v_0, h_0 w) - \tfrac{1}{2}\omega(v_0, (\mathrm{Ad}(h_0)U)v_0), \mathrm{Ad}(h_0)U\big). \end{aligned}$$

On the other hand, let us denote by $\xi = (u, d, \varphi)$, where $u \in \mathbb{C}^2$, $d \in \mathbb{C}$ and $\varphi \in (\mathfrak{s}^c)^*$, the element of $(\mathfrak{g}^c)^*$ defined by

$$\langle \xi, (w, c, U)\rangle = \omega(u, w) + cd + \langle \varphi, U\rangle.$$

Moreover, for $u, v \in \mathbb{C}^2$, we denote by $v \times u$ the element of $(\mathfrak{s}^c)^*$ defined by

$$\langle v \times u, U\rangle := \omega(u, Uv)$$

for $U \in \mathfrak{s}^c$.

Let $\xi = (u, d, \varphi) \in (\mathfrak{g}^c)^*$ and $g = (v_0, c_0, h_0) \in G^c$. Then, by using the relation

$$\big\langle \mathrm{Ad}^*(g)\xi, X\big\rangle = \big\langle \xi, \mathrm{Ad}(g^{-1})X\big\rangle$$

for each $X \in \mathfrak{g}^c$, we obtain

$$\mathrm{Ad}^*(g)\xi = \big(h_0 u - d v_0, d, \mathrm{Ad}^*(h_0)\varphi + v_0 \times (h_0 u - \tfrac{d}{2} v_0)\big)$$

By restriction, we also get the formula for the coadjoint action of G.

3 Holomorphic representations

Let us first recall the general method for constructing the holomorphic representations of G, see [22], Chapter XII. We follow the presentation of [13].

Let χ be a unitary character of K. The extension of χ to K^c is also denoted by χ. We set $K_\chi(Z,W) = \chi(\kappa(\exp W^* \exp Z))^{-1}$ for $Z, W \in \mathcal{D}$ and $J_\chi(g,Z) = \chi(\kappa(g \exp Z))$ for $g \in G$ and $Z \in \mathcal{D}$. We consider the Hilbert space $\mathcal{H}_\chi$ of holomorphic functions on $\mathcal{D}$ such that

$$\|f\|_\chi^2 := \int_{\mathcal{D}} |f(Z)|^2\, K_\chi(Z,Z)^{-1} c_\chi d\mu(Z) < +\infty$$

where the constant c_χ is defined by

$$c_\chi^{-1} = \int_{\mathcal{D}} K_\chi(Z,Z)^{-1}\, d\mu(Z)\,.$$

As we shall see, under some condition on χ, we have that c_χ is well-defined and $\mathcal{H}_\chi \neq (0)$. In that case, $\mathcal{H}_\chi$ contains the polynomials [22], p. 546. Moreover, the formula

$$\big(\pi_\chi(g)f\big)(Z) = J_\chi(g^{-1}, Z)\, f(g^{-1}\cdot Z)$$

defines a unitary representation of G on $\mathcal{H}_\chi$ which is a highest weight representation [22], p. 540.

The space $\mathcal{H}_\chi$ is a reproducing kernel Hilbert space. More precisely, if we set $e_Z(W) := K_\chi(W,Z)$ then we have we have the reproducing property $f(Z) = \langle f, e_Z\rangle_\chi$ for each $f \in \mathcal{H}_\chi$ and each $Z \in \mathcal{D}$ [22], p. 540. Here $\langle \cdot,\cdot\rangle_\chi$ denotes the inner product on $\mathcal{H}_\chi$.

Let's define χ as follows. Fix $\gamma \in \mathbb{R}$ and $m \in \mathbb{Z}$ and for $k = \big((0,0), c, \left(\begin{smallmatrix} a & 0 \\ 0 & \bar{a}\end{smallmatrix}\right)\big) \in K$ set $\chi(k) = e^{i\gamma c} a^{-m}$. Then we have the following result.

Proposition 1. *1. Let $Z = a(z,u) \in \mathcal{D}$ and $W = a(w,v) \in \mathcal{D}$. Then we have*

$$K_\chi(Z,W) = (1-u\bar{v})^{-m} \exp\left(\frac{\gamma}{4}\left(\frac{2\bar{w}z + \bar{v}z^2 + u\bar{w}^2}{1-u\bar{v}}\right)\right).$$

2. We have $\mathcal{H}_\chi \neq (0)$ if and only if $m > 3/2$. In this case, we also have $c_\chi = \frac{\gamma}{2\pi^2}(m - \frac{3}{2})$.

Proof.

1. This can be verified by a simple computation based on the $P^+K^cP^-$-decomposition.

2. By [22], Theorem XII.5.6, we have $\mathcal{H}_\chi \neq (0)$ if and only

$$I_\chi := \int_{\mathcal{D}} K_\chi(Z,Z)^{-1}\, d\mu(Z) < \infty\,.$$

Then we have to study the integral I_χ. We begin with the following remark. Let A be a symmetric positive-definite $n \times n$ real matrix. By diagonalizing A we immediately obtain that

$$\int_{\mathbb{R}^n} e^{-Ax\cdot x}\, dx = \pi^{n/2} (\mathrm{Det}\, A)^{-1/2}.$$

From this, we deduce that

$$\int_{\mathbb{C}} \exp\left(-\frac{\gamma}{4}\left(\frac{2\bar{z}z+\bar{u}z^2+u\bar{z}^2}{1-u\bar{u}}\right)\right)\,d\mu_L(z)=\frac{2\pi}{\gamma}(1-u\bar{u})^{1/2}$$

for each $u\in\mathbb{D}$. Then, taking into account the above expression for $K_\chi(Z,Z)$, we have

$$\begin{aligned} I_\chi &= \int_{\mathbb{C}\times\mathbb{D}} (1-u\bar{u})^{m-3}\exp\left(-\frac{\gamma}{4}\left(\frac{2\bar{z}z+\bar{u}z^2+u\bar{z}^2}{1-u\bar{u}}\right)\right)\,d\mu_L(z,u)\\ &= \frac{2\pi}{\gamma}\int_{\mathbb{D}}(1-u\bar{u})^{m-\frac{5}{2}}d\mu_L(u)=\frac{2\pi^2}{\gamma}\int_0^1(1-t)^{m-\frac{5}{2}}dt\,. \end{aligned}$$

Hence we have $\mathcal{H}_\chi\neq(0)$ if and only if $m>3/2$ and, in this case, we find that

$$c_\chi^{-1}=I_\chi=\frac{2\pi^2}{\gamma(m-\frac{3}{2})}\,.$$ □

Let $g=\big((z_0,\bar{z}_0),c_0,\left(\begin{smallmatrix}a & b\\ \bar{b} & \bar{a}\end{smallmatrix}\right)\big)\in G$ and $Z=a(z,u)\in\mathcal{D}$. Then, setting

$$c:=c_0+\frac{i}{4}(\bar{b}z_0-a\bar{z}_0)z-\frac{i}{4}\Big(z_0+az-\frac{au+b}{\bar{b}u+\bar{a}}(\bar{z}_0+\bar{b}z)\Big)(\bar{z}_0+\bar{b}z),$$

one can verify that $J_\chi(g,Z)=e^{i\gamma c}(\bar{b}u+\bar{a})^m$. This gives an explicit but rather complicated expression for $(\pi_\chi(g)f)(Z)$.

Now we compute the derived representation $d\pi_\chi$.

Proposition 2. *For each $X\in\mathfrak{g}^c$, $Z=a(z,u)\in\mathcal{D}$ and $f\in\mathcal{H}_\chi$, we have*

$$\big(d\pi_\chi(X)f\big)(Z)=d\chi(p_{\mathfrak{k}^c}(e^{-\operatorname{ad}Z}X))\,f(Z)-(df)_Z\big(p_{\mathfrak{p}^+}(e^{-\operatorname{ad}Z}X)\big).$$

More precisely

1. *If $X=\big((w,0),0,\left(\begin{smallmatrix}0 & v\\ 0 & 0\end{smallmatrix}\right)\big)\in\mathfrak{p}^+$ then we have $(d\pi_\chi(X)f)(Z)=-df_Z(X)$ hence*

$$d\pi_\chi(X)f=-w\partial_z f-v\partial_u f;$$

2. *If $X=\big((0,0),c,\left(\begin{smallmatrix}a & 0\\ 0 & -a\end{smallmatrix}\right)\big)\in\mathfrak{k}^c$ then we have*

$$(d\pi_\chi(X)f)(Z)=d\chi(X)f(Z)+df_Z([Z,X])$$

hence

$$d\pi_\chi(X)f=(i\gamma c-ma)f-a(z\partial_z f+2u\partial_u f)\,;$$

3. *If $X=\big((0,w),0,\left(\begin{smallmatrix}0 & 0\\ v & 0\end{smallmatrix}\right)\big)\in\mathfrak{p}^-$ then we have*

$$\begin{aligned}\big(d\pi_\chi(X)f\big)(Z)&=(d\chi\circ p_{\mathfrak{k}^c})\Big(-[Z,X]+\frac{1}{2}\big[Z,[Z,X]\big]\Big)f(Z)\\ &\quad-(df_Z\circ p_{\mathfrak{p}^+})\Big(-[Z,X]+\frac{1}{2}\big[Z,[Z,X]\big]\Big)\end{aligned}$$

hence

$$d\pi_\chi(X)f=\Big(\frac{\gamma}{4}z(2w+vz)+muv\Big)\,f+u(w+vz)\partial_z f+u^2v\partial_u f\,.$$

Proof. Taking into account the fact that $\mathfrak{p}^+$ is abelian, the first formula for $d\pi_\chi(X)$ follows from [13], Proposition 3.3 (see also [22], Proposition XII.2.1). The other formulas follows from the first formula by routine computations which are mainly based on the following fact. For each $X \in \mathfrak{p}^-$ and $Z \in \mathfrak{p}^+$ we have $[Z,[Z,[Z,X]]] = 0$ hence we get

$$e^{-\operatorname{ad} Z}(X) = X - [Z,X] + \frac{1}{2}[Z,[Z,X]]. \qquad \square$$

From this proposition we deduce the following result which will be needed later.

Proposition 3. *Let $X_1, X_2, \ldots, X_q \in \mathfrak{g}^c$. Then the operator $d\pi_\chi(X_1X_2\cdots X_q)$ is a sum of terms of the form $P_{r,s}(z,u)\partial_z^r\partial_u^s$ where $r+s \le q$ and $P_{r,s}$ is a polynomial of degree $\le 2q$.*

Proof. This result is proved by induction on q by using Proposition 2. $\square$

4 Berezin correspondence and Stratonovich-Weyl correspondence

We first review some general facts about the Berezin correspondence, the Berezin transform and the Stratonovich-Weyl correspondence.

The Berezin correspondence S_χ is defined as follows. Consider an operator (not necessarily bounded) A on $\mathcal{H}_\chi$ whose domain contains e_Z for each $Z \in \mathcal{D}$. Then the Berezin symbol of A is the function $S_\chi(A)$ defined on $\mathcal{D}$ by

$$S_\chi(A)(Z) := \frac{\langle A\, e_Z\, ,\, e_Z\rangle_\chi}{\langle e_Z\, ,\, e_Z\rangle_\chi}.$$

We can verify that each operator is determined by its Berezin symbol and that if an operator A has adjoint A^* then we have $S_\chi(A^*) = \overline{S_\chi(A)}$ [6], [7]. Moreover, for each operator A on $\mathcal{H}_\chi$ whose domain contains the coherent states e_Z for each $Z \in \mathcal{D}$ and each $g \in G$, the domain of $\pi_\chi(g^{-1})A\pi_\chi(g)$ also contains e_Z for each $Z \in \mathcal{D}$ and we have

$$S_\chi\big(\pi_\chi(g)^{-1}A\pi_\chi(g)\big)(Z) = S_\chi(A)(g\cdot Z),$$

that is, S_χ is G-equivariant, see [13] where we have also proved the following result.

Proposition 4. *1. For $g \in G$ and $Z \in \mathcal{D}$, we have*

$$S_\chi\big(\pi_\chi(g)\big)(Z) = \chi\big(\kappa(\exp Z^* g^{-1}\exp Z)^{-1}\kappa(\exp Z^*\exp Z)\big).$$

2. For $X \in \mathfrak{g}^c$ and $Z \in \mathcal{D}$, we have

$$S_\chi\big(d\pi_\chi(X)\big)(Z) = d\chi\big(p_{\mathfrak{k}^c}(\operatorname{Ad}(\zeta(\exp Z^*\exp Z)^{-1}\exp Z^*)\,X\big).$$

Consider the linear form ξ on $\mathfrak{g}^c$ defined by $\xi = -\operatorname{id}\chi$ on $\mathfrak{k}^c$ and $\xi = 0$ on $\mathfrak{p}^\pm$. Then we have $\xi(\mathfrak{g}) \subset \mathbb{R}$ and the restriction ξ_χ of ξ to $\mathfrak{g}$ is an element of $\mathfrak{g}^*$. More precisely, with the notation of Section 2 we have $\xi_\chi = (0,\gamma,\varphi_m)$ where $\varphi_m \in \mathfrak{s}^*$ is defined by $\varphi_m\left(\begin{smallmatrix} a & b \\ c & -a\end{smallmatrix}\right) = ima$.

We denote by $\mathcal{O}(\xi_\chi)$ the orbit of ξ_χ in $\mathfrak{g}^*$ for the coadjoint action of G. This orbit is said to be associated with π_χ by the Kostant-Kirillov method of orbits, see [13], [21]. Moreover, we have the following result.

Proposition 5. *[13]*

1. *For each $Z\in\mathcal{D}$, let*

$$\Psi_\chi(Z):=\mathrm{Ad}^*\big(\exp(-Z^*)\,\zeta(\exp Z^*\exp Z)\big)\,\xi_\chi\,.$$

Then, for each $X\in\mathfrak{g}^c$ and each $Z\in\mathcal{D}$, we have

$$S\big(d\pi_\chi(X)\big)(Z)=i\big\langle\Psi_\chi(Z),X\big\rangle\,.$$

2. *For each $g\in G$ and each $Z\in\mathcal{D}$, we have $\Psi_\chi(g\cdot Z)=\mathrm{Ad}^*(g)\,\Psi_\chi(Z)$.*

3. *The map Ψ_χ is a diffeomorphism from $\mathcal{D}$ onto $\mathcal{O}(\xi_\chi)$.*

We aim to make the expression of Ψ_χ more explicit. To this goal, we introduce the following notation. For $\varphi\in\mathfrak{s}^*$, let $\alpha(\varphi)$ the unique element of $\mathfrak{s}$ such that $\langle\varphi,X\rangle=\mathrm{Tr}(\alpha(\varphi)X)$ for each $X\in\mathfrak{s}$. In particular, one has $\alpha(\varphi_m)=\frac{m}{2}\left(\begin{smallmatrix} i & 0\\ 0 & -i\end{smallmatrix}\right)$. Moreover, for $u=(x,\bar{x})\in\mathbb{C}^2$ and $v=(y,\bar{y})\in\mathbb{C}^2$ we have

$$\alpha(v\times u)=\frac{i}{4}\begin{pmatrix}-x\bar{y}-\bar{x}y & 2xy\\ -2\bar{x}\bar{y} & x\bar{y}+\bar{x}y\end{pmatrix}.$$

Note also that α intertwines Ad^* and Ad. Then, by a direct calculation we obtain the following result.

Proposition 6. *Let $Z=a(z,u)\in\mathcal{D}$. Let us denote by $\varphi(z,u)$ the element of $\mathfrak{s}^*$ defined by*

$$\begin{aligned}\alpha\big(\varphi(z,u)\big)&=\frac{mi}{2}(1-u\bar{u})^{-1}\begin{pmatrix}1+u\bar{u} & -2u\\ 2\bar{u} & -(1+u\bar{u})\end{pmatrix}\\ &\quad-\frac{\gamma i}{4}(1-u\bar{u})^{-2}\begin{pmatrix}-|z+u\bar{z}|^2 & (z+u\bar{z})^2\\ -(\bar{z}+\bar{u}z)^2 & |z+u\bar{z}|^2\end{pmatrix}.\end{aligned}$$

Then we have

$$\Psi_\chi(Z)=\big(-\gamma(1-u\bar{u})^{-1}(z+u\bar{z},\bar{z}+\bar{u}z),\gamma,\varphi(z,u)\big)\,.$$

Let us recall the construction of the Stratonovich-Weyl correspondence [12], [13], [17]. Denote by $\mathcal{L}_2(\mathcal{H}_\chi)$ the space of all Hilbert-Schmidt operators on $\mathcal{H}_\chi$ and by μ_χ the G-invariant measure on $\mathcal{D}$ defined by $d\mu_\chi(Z)=c_\chi d\mu(Z)$. Then the map S_χ is a bounded operator from $\mathcal{L}_2(\mathcal{H}_\chi)$ into $L^2(\mathcal{D},\mu_\chi)$ which is one-to-one and has dense range [25], [28].

The Berezin transform is the operator on $L^2(\mathcal{D},\mu_\chi)$ defined by $B_\chi:=S_\chi S_\chi^*$. We can easily verify that we have

$$B_\chi F(Z)=\int_{\mathcal{D}}F(W)\,\frac{|\langle e_Z,e_W\rangle|_\chi^2}{\langle e_Z,e_Z\rangle_\chi\langle e_W,e_W\rangle_\chi}\,d\mu_\chi(W)\qquad(1)$$

(see [6], [28], [29] for instance).

Let ρ be the left-regular representation of G on $L^2(\mathcal{D},\mu_\chi)$. As a consequence of the equivariance property for S_χ, we see that B_χ commute with $\rho(g)$ for each $g\in G$.

Let us consider the polar decomposition of S_χ:

$$S_\chi=(S_\chi S_\chi^*)^{1/2}W_\chi=B_\chi^{1/2}W_\chi$$

where $W_\chi:=B_\chi^{-1/2}S_\chi$ is a unitary operator from $\mathcal{L}_2(\mathcal{H}_\chi)$ onto $L^2(\mathcal{D},\mu_\chi)$. We immediately obtain the following proposition. Note that, by 2. of Proposition 5, the measure $\mu_0:=(\Psi_\chi^{-1})^*(\mu_\chi)$ is a G-invariant measure on $\mathcal{O}(\xi_\chi)$.

Proposition 7. *1) The map $W_\chi\colon \mathcal{L}_2(\mathcal{H}_\chi)\to L^2(\mathcal{D},\mu_\chi)$ is a Stratonovich-Weyl correspondence for the triple $(G,\pi_\chi,\mathcal{D})$, that is, we have*

1. $W_\chi(A^*)=\overline{W_\chi(A)}$;
2. $W_\chi(\pi_\chi(g)\,A\,\pi_\chi(g)^{-1})(Z)=W_\chi(A)(g^{-1}\cdot Z)$;
3. W_χ *is unitary.*

2) Similarly, the map $\mathcal{W}_\chi\colon \mathcal{L}_2(\mathcal{H}_\chi)\to L^2(\mathcal{O}(\xi_\chi),\mu_0)$ defined by

$$\mathcal{W}_\chi(A)=W_\chi(A)\circ\Psi_\chi^{-1}$$

is a Stratonovich-Weyl correspondence for the triple $(G,\pi_\chi,\mathcal{O}(\xi_\chi))$.

Here we have relaxed the requirement of Definition 1 that W_χ maps the identity operator I to the constant function 1 which is not adapted to the present setting (here I is not Hilbert-Schmidt). However, this requirement should hold in some generalized sense, see for instance [19].

5 Extension of the Berezin transform

In this section, we show how to extend the Berezin transform to a class of functions which contains $S_\chi(d\pi_\chi(X))$ for each $X\in\mathfrak{g}^c$, in order to define and study $W_\chi(d\pi_\chi(X))$. This question was investigated in [13] but only in the case of a reductive Lie group. Here we adapt the method of [13], Section 6 to the case of the Jacobi group.

We use the following notation. If L is a Lie group and X is an element of the Lie algebra of L then we denote by X^+ the corresponding right invariant vector field on L, that is, $X^+(h)=\frac{d}{dt}(\exp tX)h|_{t=0}$ for $h\in L$. By differentiating the multiplication map from $P^+\times K^c\times P^-$ onto $P^+K^cP^-$, we can easily prove the following result.

Lemma 1. *[10], [13] Let $X\in\mathfrak{g}^c$ and $g=z\,k\,y$ where $z\in P^+$, $k\in K^c$ and $y\in P^-$. We have*

1. $d\zeta_g\big(X^+(g)\big)=\big(\mathrm{Ad}(z)\,p_{\mathfrak{p}^+}(\mathrm{Ad}(z^{-1})\,X)\big)^+(z)$.
2. $d\kappa_g\big(X^+(g)\big)=\big(p_{\mathfrak{k}^c}(\mathrm{Ad}(z^{-1})\,X)\big)^+(k)$.

3. $d\eta_g\big(X^+(g)\big) = \big(\mathrm{Ad}(k^{-1})\, p_{\mathfrak{p}^-}(\mathrm{Ad}(z^{-1})\, X)\big)^+(y)$.

For Z, $W \in \mathcal{D}$, we set $l_Z(W) := \log \eta(\exp Z^* \exp W) \in \mathfrak{p}^-$.

Lemma 2. *1. For each Z, $W \in \mathcal{D}$ and $V \in \mathfrak{p}^+$, we have*

$$\frac{d}{dt}\, e_Z(W+tV)\big|_{t=0} = -e_Z(W)\,(d\chi\circ p_{\mathfrak{k}^c})\Big([l_Z(W),V]+\frac{1}{2}\big[l_Z(W),[l_Z(W),V]\big]\Big).$$

2. For each Z, $W \in \mathcal{D}$ and $V \in \mathfrak{p}^+$, we have

$$\frac{d}{dt}\, l_Z(W+tV)\big|_{t=0} = p_{\mathfrak{p}^-}\Big([l_Z(W),V]+\frac{1}{2}\big[l_Z(W),[l_Z(W),V]\big]\Big).$$

3. Let r and s be non-negative integers such that $r+s \le q$. Let $Z \in \mathcal{D}$. Then the function $(\partial_z^r \partial_u^s e_Z)(W)$ is of the form $e_Z(W)Q(l_Z(W))$ where Q is a polynomial on $\mathfrak{p}^-$ of degree $\le 2q$.

4. Let r and s be non-negative integers such that $r+s \le q$. Then $S_\chi(\partial_z^r \partial_u^s)$ is of the form $(1-u\bar{u})^{-2q} P(z,\bar{z},u,\bar{u})$ where P is a polynomial whose degree in $(z,\bar{z})$ is $\le 2q$.

5. For each $X_1, X_2, \ldots, X_q \in \mathfrak{g}^c$, the function $S_\chi(d\pi_\chi(X_1X_2\cdots X_q))(Z)$ is of the form $(1-u\bar{u})^{-2q} P(z,\bar{z},u,\bar{u})$ where P is a polynomial whose degree in $(z,\bar{z})$ is $\le 4q$.

Proof. Following the same lines as in the proof of Lemma 4.1 in [12], we use 2. of Lemma 1. For Z, $W \in \mathcal{D}$ and $V \in \mathfrak{p}^+$, we have

$$\begin{aligned}
\frac{d}{dt}\, e_Z(W+tV)\big|_{t=0} &= \frac{d}{dt}\, \chi^{-1}\left(\kappa(\exp Z^* \exp W \exp tV)\right)\big|_{t=0} \\
&= d\chi^{-1}_{\kappa(\exp Z^* \exp W)} d\kappa_{\exp Z^* \exp W}\Big(\big(\mathrm{Ad}(\exp Z^* \exp W)\, V\big)^+(\exp Z^* \exp W)\Big) \\
&= -\chi^{-1}\big(\kappa(\exp Z^* \exp W)\big)\, d\chi\Big(p_{\mathfrak{k}^c}\Big(\mathrm{Ad}\big(\kappa(\exp Z^* \exp W)\eta(\exp Z^* \exp W)\big)V\Big)\Big) \\
&= -e_Z(W)(d\chi\circ p_{\mathfrak{k}^c})\Big(\mathrm{Ad}\big(\eta(\exp Z^* \exp W)\big)V\Big).
\end{aligned}$$

But here we have $[\mathfrak{p}^-,[\mathfrak{p}^-,[\mathfrak{p}^-,\mathfrak{p}^+]]] = (0)$. This implies that

$$p_{\mathfrak{k}^c}\Big(\mathrm{Ad}\big(\eta(\exp Z^* \exp W)\big)V\Big) = p_{\mathfrak{k}^c}\Big([l_Z(W),V]+\frac{1}{2}\big[l_Z(W),[l_Z(W),V]\big]\Big).$$

Hence Statement 1. is proved. Similarly, we prove 2. From 1. and 2., we deduce 3. by induction on q.

Now, let $Z = a(z,u) \in \mathcal{D}$. Then we have

$$l_Z(Z) = \Big(\big(0, -(1-u\bar{u})^{-1}(\bar{z}+\bar{u}z)\big), 0, \left(\begin{smallmatrix} 0 & 0 \\ -(1-u\bar{u})^{-1}\bar{u} & 0 \end{smallmatrix}\right)\Big)$$

and, by 3., $S_\chi(\partial_z^r \partial_u^s)(Z) = Q(l_Z(Z))$ is a linear combination of terms of the form

$$\big((1-u\bar{u})^{-1}(\bar{z}+\bar{u}z)\big)^i \big((1-u\bar{u})^{-1}\bar{u}\big)^j$$

for $i+j \le 2q$. This gives 4. Finally, 5. follows from 4. and Proposition 3. $\square$

We are now in position to extend the Berezin transform. The following proposition is analogous to Proposition 6.3 of [13] (see also Proposition 4.1 of [12]).

Proposition 8. *If $m > 2q + \frac{3}{2}$ then for each $X_1, X_2, \ldots, X_q \in \mathfrak{g}^c$, the Berezin transform of $S_\chi\big(d\pi_\chi(X_1X_2\cdots X_q)\big)$ is well-defined.*

Proof. First recall the equality $e_{g\cdot Z} = \overline{\chi(\kappa(g\exp Z))}\pi_\chi(g)e_Z$ for each $Z \in \mathcal{D}$ and each $g \in G$ [13]. Then, by performing the change of variables $W \to g_Z \cdot W$ in equation (1), we obtain

$$(B_\chi F)(Z) = \int_{\mathcal{D}} F(g_Z \cdot W)\langle e_W, e_W\rangle_\chi^{-1}\, d\mu_\chi(W).$$

Now we take $F = S_\chi\big(d\pi_\chi(X_1X_2\cdots X_q)\big)$. Fixing $Z \in \mathcal{D}$, we set $Y_k := \mathrm{Ad}(g_Z^{-1})X_k$ for $k = 1, 2, \ldots, q$. Then, by using the G-invariance of S_χ, we get

$$F(g_Z \cdot W) = S_\chi(d\pi_\chi(Y_1Y_2\cdots Y_q))(W)$$

for each $W \in \mathcal{D}$. Thus, by 1. of Proposition 1 and 5. of Lemma 2, we see that the proposition will be established if we prove that under the condition that $m > 2q + \frac{3}{2}$ the integral

$$I := \int_{\mathbb{C}\times\mathbb{D}} P(z, \bar z, u, \bar u)(1 - u\bar u)^{-2q+m-3}\exp\left(-\frac{\gamma}{4}\left(\frac{2\bar z z + \bar u z^2 + u\bar z^2}{1 - u\bar u}\right)\right) d\mu_L(z, u)$$

converges for each polynomial $P(z, \bar z, u, \bar u)$ whose degree in $(z, \bar z)$ is $\leq 2q$.

We set $z = x + iy$ with $x, y \in \mathbb{R}$ and $u = a + ib$ with $a, b \in \mathbb{R}$. Then we have

$$2z\bar z + \bar u z^2 + u\bar z^2 = 2\big((1 + a)x^2 + (1 - a)y^2 + 2bxy\big).$$

For $u \neq 0$, this quadratic form can be reduced by means of an orthonormal change of variables of the form $z \to vz$ ($|v| = 1$). Under this change of variables, the integral I becomes

$$I = \int_{\mathbb{C}\times\mathbb{D}} P(vz, \overline{vz}, u, \bar u)(1 - u\bar u)^{-2q+m-3} \times \exp\left(-\frac{\gamma}{2}\left(\frac{(1 + |u|)x^2 + (1 - |u|)y^2}{1 - u\bar u}\right)\right) d\mu_L(z, u)\,.$$

Now, we make the last change of variables

$$x' = (1 - |u|)^{-1/2}x, \quad y' = (1 + |u|)^{-1/2}y$$

and we obtain

$$I = \int_{\mathbb{C}\times\mathbb{D}} P(v(x'\sqrt{1 - |u|} + iy'\sqrt{1 + |u|}), \bar v(x'\sqrt{1 - |u|} - iy'\sqrt{1 + |u|}), u, \bar u) \times (1 - u\bar u)^{-2q+m-\frac{5}{2}}e^{-\frac{\gamma}{2}|z'|^2}\, d\mu_L(z', u).$$

Finally we see that, under the condition that $m > 2q + \frac{3}{2}$, we have

$$\int_{\mathbb{D}} (1 - u\bar u)^{-2q+m-\frac{5}{2}}\, d\mu_L(u) < +\infty$$

hence I converges. This ends the proof. □

6 Stratonovich-Weyl symbols of derived representation operators

In this section, we assume that $m > \frac{7}{2}$. Then, by Proposition 8, the Berezin transform of $S_\chi(d\pi_\chi(X))$ is well-defined for each $X \in \mathfrak{g}^c$. As we shall see, this fact can be used to define $W_\chi(d\pi_\chi(X))$ for $X \in \mathfrak{g}^c$. The first step is to introduce a vector space of functions on $\mathcal{D}$ which is stable under B_χ and contains $S_\chi(d\pi_\chi(X))$ for each $X \in \mathfrak{g}^c$.

Note that, for each $X \in \mathfrak{g}^c$ and $Z \in \mathcal{D}$, we have $S_\chi(d\pi_\chi(X))(Z) = i\xi(\mathrm{Ad}(g_Z^{-1})X)$ by Proposition 5. Then, we introduce the space $\mathcal{S}$ generated by the functions $Z \to \phi_0(\mathrm{Ad}(g_Z^{-1})X)$ where $X \in \mathfrak{g}^c$ and ϕ_0 is an element of $(\mathfrak{g}^c)^*$ which is $\mathrm{Ad}^*(K)$-invariant. The following lemma can be easily verified, see [14].

Lemma 3. *Let φ_0 denote the element of $(\mathfrak{s}^c)^*$ defined by $\langle \varphi_0, \left(\begin{smallmatrix} a & b \\ c & -a \end{smallmatrix}\right)\rangle = a$. Then the elements of $(\mathfrak{g}^c)^*$ which are fixed by K are the elements of the form $(0, d, \lambda\varphi_0)$ where $d, \lambda \in \mathbb{C}$.*

Then we have the following result.

Proposition 9. *Let $\phi\colon \mathcal{D} \times \mathfrak{g}^c \to \mathbb{C}$ be a function such that*
(i) For each $Z \in \mathcal{D}$, the map $X \to \phi(Z, X)$ is linear;
(ii) For each $X \in \mathfrak{g}^c$, $g \in G$ and $Z \in \mathcal{D}$, we have $\phi(g \cdot Z, X) = \phi(Z, \mathrm{Ad}(g^{-1})X)$.
Then

1. *The element ϕ_0 of $(\mathfrak{g}^c)^*$ defined by $\phi_0(X) := \phi(0, X)$ is fixed by K;*

2. *For each $X \in \mathfrak{g}^c$ and $Z \in \mathcal{D}$, we have*

$$\begin{aligned}\phi(Z, X) &= \phi_0(\mathrm{Ad}(g_Z^{-1})X) = \phi_0\left(\mathrm{Ad}(\zeta(\exp Z^* \exp Z)^{-1} \exp Z^*)X\right)\\ &= (\phi_0 \circ p_{\mathfrak{k}^c})\left(\mathrm{Ad}(\zeta(\exp Z^* \exp Z)^{-1} \exp Z^*)X\right).\end{aligned}$$

   In particular, for each $X \in \mathfrak{g}^c$, the function $\phi(\cdot, X)$ is in $\mathcal{S}$.

3. *For each $X \in \mathfrak{g}^c$, the Berezin transform of the function $\phi(\cdot, X)$ is well-defined and there exists a function $\psi\colon \mathcal{D} \times \mathfrak{g}^c \to \mathbb{C}$ satisfying (i) and (ii) such that, for each $X \in \mathfrak{g}^c$, one has $\psi(\cdot, X) = B_\chi(\phi(\cdot, X))$.*

Proof. 1. This follows from (ii).

2. By (ii) again, we have $\phi(Z, X) = \phi_0(\mathrm{Ad}(g_Z^{-1})X)$ for each $X \in \mathfrak{g}^c$ and $Z \in \mathcal{D}$. Now, by [14], there exists $k_Z \in K$ such that $g_Z = \exp(-Z^*)\zeta(\exp Z^* \exp Z)k_Z^{-1}$. Then we have

$$\begin{aligned}\phi(Z, X) &= \phi_0\left(\mathrm{Ad}(k_Z\zeta(\exp Z^* \exp Z)^{-1} \exp Z^*)X\right)\\ &= \phi_0\left(\mathrm{Ad}(\zeta(\exp Z^* \exp Z)^{-1} \exp Z^*)X\right)\end{aligned}$$

and, since $\phi_0|_{\mathfrak{p}^\pm} = 0$ by Lemma 3, we also have

$$\phi(Z, X) = (\phi_0 \circ p_{\mathfrak{k}^c})\left(\mathrm{Ad}(\zeta(\exp Z^* \exp Z)^{-1} \exp Z^*)X\right).$$

3. We prove the first assertion by the same method as in the proof of Proposition 8. The second assertion immediately follows from the fact that B_χ commutes to the $\rho(g)$, $g \in G$ (see Section 4). □

Now we aim to compute the Berezin transform of a function $\phi(\cdot, X)$ as before. We need the following lemma.

Lemma 4. *For each $u \in \mathbb{D}$, we have*

$$I_1(u) = \int_{\mathbb{C}} \exp\left(-\frac{\gamma}{4}\left(2\bar{z}z + \bar{u}z^2 + u\bar{z}^2\right)\right) d\mu_L(z) = \frac{2\pi}{\gamma}(1-u\bar{u})^{-1/2};$$
$$I_2(u) = \int_{\mathbb{C}} z^2 \exp\left(-\frac{\gamma}{4}\left(2\bar{z}z + \bar{u}z^2 + u\bar{z}^2\right)\right) d\mu_L(z) = -\frac{4\pi}{\gamma^2}u(1-u\bar{u})^{-3/2};$$
$$I_3(u) = \int_{\mathbb{C}} \bar{z}^2 \exp\left(-\frac{\gamma}{4}\left(2\bar{z}z + \bar{u}z^2 + u\bar{z}^2\right)\right) d\mu_L(z) = -\frac{4\pi}{\gamma^2}\bar{u}(1-u\bar{u})^{-3/2};$$
$$I_4(u) = \int_{\mathbb{C}} z\bar{z} \exp\left(-\frac{\gamma}{4}\left(2\bar{z}z + \bar{u}z^2 + u\bar{z}^2\right)\right) d\mu_L(z) = \frac{4\pi}{\gamma^2}(1-u\bar{u})^{-3/2}.$$

Proof. The integral $I_1(u)$ has been computed in the proof of Proposition 1. By taking the derivative of $I_1(u)$ with respect to u, we obtain the desired expression of $I_2(u)$ and, by conjugation, the expression of $I_3(u)$. Finally, integrating by parts, we have

$$I_1(u) = -\int_{\mathbb{C}} z\partial_z\left(\exp\left(-\frac{\gamma}{4}\left(2\bar{z}z + \bar{u}z^2 + u\bar{z}^2\right)\right)\right) d\mu_L(z) = \frac{\gamma}{2}\big(I_4(u) + \bar{u}I_2(u)\big)$$

and we get the expression of $I_4(u)$. □

We introduce the following basis of $\mathfrak{g}^c$:

$$X_1 = \big((1,0),0,0\big); \qquad Y_1 = \big((0,0),0,\left(\begin{smallmatrix}1 & 0\\ 0 & 0\end{smallmatrix}\right)\big); \qquad H_1 = \big((0,0),1,0\big);$$
$$X_2 = \big((0,1),0,0\big); \qquad Y_2 = \big((0,0),0,\left(\begin{smallmatrix}0 & 0\\ 0 & 1\end{smallmatrix}\right)\big); \qquad H_2 = \big((0,0),0,\left(\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}\right)\big).$$

Also, we denote by ϕ^1 and ϕ^2 the elements of $\mathcal{S}$ defined by $\phi^1_0 = (0,1,0)$ and $\phi^2_0 = (0,0,\varphi_0)$.

Proposition 10. *Let $\mu := \frac{1}{\gamma}i\frac{2m-3}{2m-5}$ and $\nu := \frac{2m-1}{2m-5}$. Let $\phi \in \mathcal{S}$ be defined by $\phi_0 = (0,d,\lambda\varphi_0)$ with $d,\lambda \in \mathbb{C}$. Let $\psi \in \mathcal{S}$ such that $\psi(\cdot,X) = B_\chi(\phi(\cdot,X))$ for each $X \in \mathfrak{g}^c$. Then we have $\psi_0 = (0,d,d\mu + \lambda\nu)$.*

Proof. We have just to compute the Berezin transforms $\psi^1(\cdot,X)$ and $\psi^2(\cdot,X)$ of $\phi^1(\cdot,X)$ and $\phi^2(\cdot,X)$.

We can write $\psi^1_0 = (0,d_1,\lambda_1\varphi_0)$ with $d_1,\lambda_1 \in \mathbb{C}$. For each $Z \in \mathcal{D}$, we have $\mathrm{Ad}(g_Z^{-1})H_1 = H_1$ hence $\phi^1(Z,H_1) = \phi^1_0(\mathrm{Ad}(g_Z^{-1})H_1) = 1$. This implies that

$$\psi^1_0(H_1) = \int_{\mathcal{D}} \langle e_Z, e_Z\rangle_\chi^{-1} d\mu_\chi(Z) = 1.$$

Consequently, we find $d_1 = 1$. On the other hand, we can verify that, for each $Z = a(z,u) \in \mathcal{D}$, we have

$$\phi^1(Z,H_2) = \phi^1_0\big(\mathrm{Ad}(g_Z^{-1})H_2\big) = \frac{i}{2}(1-u\bar{u})^{-2}\big(\bar{u}z^2 + (1+u\bar{u})z\bar{z} + u\bar{z}^2\big).$$

Then we get

$$\begin{aligned}\psi_0^1(H_2) &= \frac{i}{2}c_\chi \int_{\mathbb{C}\times\mathbb{D}} (1-u\bar{u})^{m-5}\big(\bar{u}z^2+(1+u\bar{u})z\bar{z}+u\bar{z}^2\big)\\ &\quad\times \exp\left(-\frac{\gamma}{4}\left(\frac{2\bar{z}z+\bar{u}z^2+u\bar{z}^2}{1-u\bar{u}}\right)\right) d\mu_L(z,u)\\ &= \frac{i}{2}c_\chi \int_{\mathbb{C}\times\mathbb{D}} (1-u\bar{u})^{m-3}\big(\bar{u}z^2+(1+u\bar{u})z\bar{z}+u\bar{z}^2\big)\\ &\quad\times \exp\left(-\frac{\gamma}{4}\left(2\bar{z}z+\bar{u}z^2+u\bar{z}^2\right)\right) d\mu_L(z,u)\end{aligned}$$

and, using Lemma 4, we obtain

$$\psi_0^1(H_2) = \frac{i}{2}c_\chi\frac{4\pi}{\gamma^2}\int_{\mathbb{D}}(1-u\bar{u})^{m-\frac{7}{2}}d\mu_L(u) = \frac{i}{2}c_\chi\frac{4\pi^2}{\gamma^2}\frac{1}{m-\frac{5}{2}}\,.$$

Thus, taking into account the value of c_χ (see Proposition 1), we can conclude that $\psi_0^1(H_2)=\mu$ hence $\lambda_1=\mu$.

Similarly, we write $\psi_0^2=(0,d_2,\lambda_2\varphi_0)$ with $d_2,\lambda_2\in\mathbb{C}$ and, since we have

$$\phi^2(Z,H_1)=\phi_0^2(\mathrm{Ad}(g_Z^{-1})H_1)=0$$

and

$$\phi^2(Z,H_2)=\phi_0^2(\mathrm{Ad}(g_Z^{-1})H_2)=(1-u\bar{u})^{-1}(1+u\bar{u})$$

we obtain $d_2=\psi_0^2(H_1)=0$ and $\lambda_2=\psi_0^2(H_2)=\nu$. □

Now we show that $B_\chi\colon\mathcal{S}\to\mathcal{S}$ can be diagonalized and has positive eigenvalues.

Lemma 5. *The functions $\phi^1(\cdot,X_1)$, $\phi^1(\cdot,X_2)$, $\phi^1(\cdot,H_1)$, $\phi^1(\cdot,Y_1)$,$\phi^1(\cdot,Y_2)$, $\phi^1(\cdot,H_2)$, $\phi^2(\cdot,Y_1)$, $\phi^2(\cdot,Y_2)$, $\phi^2(\cdot,H_2)$ form a basis $\mathcal{B}_0$ of $\mathcal{S}$.*

Proof. Note that $\phi^2(\cdot,X_1)=\phi^2(\cdot,X_2)=\phi^2(\cdot,H_1)=0$. Moreover, we can easily verify that the equality $\phi^1(\cdot,X)=\phi^2(\cdot,Y)$ where $X\in\mathfrak{g}^c$ and $Y\in\mathrm{Span}_{\mathbb{C}}\{Y_1,Y_2,H_2\}$ implies $X=Y=0$. The result follows. □

Let us introduce the functions

$$\begin{aligned}s_1 &:= (1-\nu)\phi^1(\cdot,Y_1)+\mu\phi^2(\cdot,Y_1)\,;\\ s_2 &:= (1-\nu)\phi^1(\cdot,Y_2)+\mu\phi^2(\cdot,Y_2)\,;\\ s_3 &:= (1-\nu)\phi^1(\cdot,H_2)+\mu\phi^2(\cdot,H_2)\,.\end{aligned}$$

By combining the previous lemma with Proposition 10 we immediately obtain the following result.

Lemma 6. *The functions $\phi^1(\cdot,X_1)$, $\phi^1(\cdot,X_2)$, $\phi^1(\cdot,H_1)$, s_1, s_2, s_3, $\phi^2(\cdot,Y_1)$, $\phi^2(\cdot,Y_2)$, $\phi^2(\cdot,H_2)$ form a basis of $\mathcal{S}$ consisting in eigenvectors of B_χ. With respect* *to this basis,* $B_\chi|_{\mathcal{S}}$ *has matrix* $\mathrm{Diag}(1,1,1,1,1,1,\nu,\nu,\nu)$*. Moreover, the matrix of* $B_\chi|_{\mathcal{S}}$ *with respect to* $\mathcal{B}_0$ *is the* 9×9 *matrix*

$$\begin{pmatrix} I_3 & O & O \\ O & I_3 & O \\ O & \mu I_3 & \nu I_3 \end{pmatrix}$$

where I_3 *denotes the* 3×3 *identity matrix.*

Recall that for each $X\in\mathfrak{g}^c$, we have $S_\chi(d\pi_\chi(X))\in\mathcal{S}$. Then we see that the expression $W_\chi(d\pi_\chi(X))=B_\chi^{-1/2}(S_\chi(d\pi_\chi(X)))$ is well-defined. More precisely, we have the following result.

Proposition 11. *For each* $X\in\mathrm{Span}_{\mathbb{C}}\{X_1,X_2,H_1\}$*, we have*

$$W_\chi\big(d\pi_\chi(X)\big)=S_\chi\big(d\pi_\chi(X)\big)\,.$$

For each $X\in\mathrm{Span}_{\mathbb{C}}\{Y_1,Y_2,H_2\}$*, we have*

$$W_\chi\big(d\pi_\chi(X)\big)=S_\chi\big(d\pi_\chi(X)\big)+(1-\nu^{-1/2})\left(\frac{i\gamma\mu}{1-\nu}+m\right)\phi^2(\cdot,X)\,.$$

Proof. Recall that for each $X\in\mathfrak{g}^c$ we have

$$S_\chi(d\pi_\chi(X))(Z)=d\chi(\mathrm{Ad}(g_Z^{-1})X)=i\gamma\phi^1(Z,X)-m\phi^2(Z,X).$$

From Lemma 6, we deduce that the matrix of $B_\chi^{-1/2}$ with respect to $\mathcal{B}_0$ is

$$\begin{pmatrix} I_3 & O & O \\ O & I_3 & O \\ O & -\frac{\mu\nu^{-1/2}}{1+\nu^{1/2}}I_3 & \nu^{-1/2}I_3 \end{pmatrix}.$$

The result then follows. □

Acknowledgements

I thank the referee for a careful reading of the manuscript and many valuable comments.

References

[1] S.T. Ali, M. Englis: Quantization methods: a guide for physicists and analysts. *Rev. Math. Phys.* 17 (4) (2005) 391–490.

[2] J. Arazy, H. Upmeier: *Invariant symbolic calculi and eigenvalues of invariant operators on symmetric domains.* Function spaces, interpolation theory and related topics (Lund, 2000) 151–211. De Gruyter, Berlin (2002).

[3] J. Arazy, H. Upmeier: Weyl Calculus for Complex and Real Symmetric Domains, Harmonic analysis on complex homogeneous domains and Lie groups (Rome, 2001). *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl.* 13 (3–4) (2002) 165–181.

[4] S. Berceanu: A holomorphic representation of the Jacobi algebra. *Rev. Math. Phys.* 18 (2006) 163–199.

[5] S. Berceanu, A. Gheorghe: On the geometry of Siegel-Jacobi domains. *Int. J. Geom. Methods Mod. Phys.* 8 (2011) 1783–1798.

[6] F.A. Berezin: Quantization. *Math. USSR Izv.* 8 (5) (1974) 1109–1165.

[7] F.A. Berezin: Quantization in complex symmetric domains. *Math. USSR Izv.* 9 (2) (1975) 341–379.

[8] R. Berndt, S. Böcherer: Jacobi forms and discrete series representations of the Jacobi group. *Math. Z.* 204 (1990) 13–44.

[9] R. Berndt, R. Schmidt: *Elements of the representation theory of the Jacobi group,* Progress in Mathematics 163. Birkhäuser Verlag, Basel (1998).

[10] B. Cahen: Berezin quantization for discrete series. *Beiträge Algebra Geom.* 51 (2010) 301–311.

[11] B. Cahen: Stratonovich-Weyl correspondence for compact semisimple Lie groups. *Rend. Circ. Mat. Palermo* 59 (2010) 331–354.

[12] B. Cahen: Stratonovich-Weyl correspondence for discrete series representations. *Arch. Math. (Brno)* 47 (2011) 41–58.

[13] B. Cahen: Berezin Quantization and Holomorphic Representations. *Rend. Sem. Mat. Univ. Padova* 129 (2013) 277–297.

[14] B. Cahen: Global Parametrization of Scalar Holomorphic Coadjoint Orbits of a Quasi-Hermitian Lie Group. *Acta Univ. Palacki. Olomuc., Fac. rer. nat., Mathematica* 52 (2013) 35–48.

[15] J.F. Cariñena, J.M. Gracia-Bondìa, J.C. Vàrilly: Relativistic quantum kinematics in the Moyal representation. *J. Phys. A: Math. Gen.* 23 (1990) 901–933.

[16] M. Davidson, G. Òlafsson, G. Zhang: Laplace and Segal-Bargmann transforms on Hermitian symmetric spaces and orthogonal polynomials. *J. Funct. Anal.* 204 (2003) 157–195.

[17] H. Figueroa, J.M. Gracia-Bondìa, J.C. Vàrilly: Moyal quantization with compact symmetry groups and noncommutative analysis. *J. Math. Phys.* 31 (1990) 2664–2671.

[18] B. Folland: *Harmonic Analysis in Phase Space.* Princeton Univ. Press (1989).

[19] J.M. Gracia-Bondìa: *Generalized Moyal quantization on homogeneous symplectic spaces,* Deformation theory and quantum groups with applications to mathematical physics (Amherst, MA, 1990), 93–114, Contemp. Math., 134. Amer. Math. Soc., Providence, RI (1992).

[20] J.M. Gracia-Bondìa, J.C. Vàrilly: The Moyal Representation for Spin. *Ann. Physics* 190 (1989) 107–148.

[21] A.A. Kirillov: *Lectures on the Orbit Method,* Graduate Studies in Mathematics, Vol. 64. American Mathematical Society, Providence, Rhode Island (2004).

[22] K-H. Neeb: *Holomorphy and Convexity in Lie Theory,* de Gruyter Expositions in Mathematics, Vol. 28. Walter de Gruyter, Berlin, New-York (2000).

[23] T. Nomura: Berezin Transforms and Group representations. *J. Lie Theory* 8 (1998) 433–440.

[24] B. Ørsted, G. Zhang: Weyl Quantization and Tensor Products of Fock and Bergman Spaces. *Indiana Univ. Math. J.* 43 (2) (1994) 551–583.

[25] J. Peetre, G. Zhang: A weighted Plancherel formula III. The case of a hyperbolic matrix ball. *Collect. Math.* 43 (1992) 273–301.

[26] I. Satake: *Algebraic structures of symmetric domains.* Iwanami Sho-ten, Tokyo and Princeton Univ. Press, Princeton, NJ (1971).

[27] R.L. Stratonovich: On distributions in representation space. *Soviet Physics. JETP* 4 (1957) 891–898.

[28] A. Unterberger, H. Upmeier: Berezin transform and invariant differential operators. *Commun. Math. Phys.* 164 (3) (1994) 563–597.

[29] G. Zhang: Berezin transform on compact Hermitian symmetric spaces. *Manuscripta Math.* 97 (1998) 371–388.

Author's address:
Université de Lorraine, Site de Metz, UFR-MIM, Département de mathématiques, Bâtiment A, Ile du Saulcy, CS 50128, F-57045, Metz cedex 01, France.
E-mail: benjamin.cahen@univ-lorraine.fr

Received: 2 December, 2013
Accepted for publication: 2 March, 2014
Communicated by: Ilka Agricola

Communications in Mathematics 22 (2014) 49–55

A note on the number of S-Diophantine quadruples

Florian Luca, Volker Ziegler

Abstract. Let $(a_1, \dots, a_m)$ be an m-tuple of positive, pairwise distinct integers. If for all $1 \le i < j \le m$ the prime divisors of $a_i a_j + 1$ come from the same fixed set S, then we call the m-tuple S-Diophantine. In this note we estimate the number of S-Diophantine quadruples in terms of $|S| = r$.

1 Introduction

There is a vast amount of papers concerning the problem of determining the number of prime divisors of products of the form

$$\prod_{a \in A,\, b \in B} (a+b) \qquad \text{and} \qquad \prod_{a \in A,\, b \in B} (ab+1),$$

where A and B are finite sets of positive integers. In particular, the first product has been studied, first by Erdős and Turán [4] and their investigations were continued in a series of papers by Sárközy and Stewart (see e.g. [12], [13]). The second product was studied e.g. by Győry, Sárközy and Stewart [8], Sárközy and Stewart [14], and others.

In their paper [8], Győry, Sárközy and Stewart conjectured that the largest prime factor of

$$(ab+1)(ac+1)(bc+1)\,, \quad 0 < a < b < c\,,$$

goes to infinity as c does. This conjecture has been proved by Corvaja and Zannier [3] and Hernandez and Luca [9], independently. Due to the application of the Subspace theorem their results stay ineffective. The best approach to estimate the growth rate of the largest prime factor of $(ab+1)(ac+1)(bc+1)$ is due to Luca [10], who proved that for every fixed finite set of primes S, there exist ineffective constants C_S and C'_S such that

$$((bc+1)(ac+1))_{\bar{S}} > \exp\left(C_S \frac{\log c}{\log\log c}\right)$$

2010 MSC: 11D45, 11N32
Key words: Diophantine equations, S-unit equations, S-Diophantine tuples

whenever $a < b < c$ with $c > C'_S$, where $(\cdot)_{\bar{S}}$ denotes the S-free part.

In case of quadruples effective results are known. For example, Stewart and Tijdeman [15], proved that the largest prime factor of

$$\prod_{a,b\in A,\, a\neq b} (ab+1)$$

with $|A| \geq 4$, is at least $C \log\log \max A$, where C is an effective computable constant.

Let S be a fixed, finite set of primes. In view of classical Diophantine m-tuples we call an m-tuple $(a_1, \dots, a_m)$ of positive, pairwise distinct, integers S-Diophantine if for all $1 \leq i < j \leq m$ the set of prime divisors of $a_i a_j + 1$ is contained in S. The results of Corvaja, Zannier [3] and Hernandez, Luca [9] yield the finiteness of S-Diophantine triples for fixed S. Although we are able to estimate the number of S-Diophantine triples due to a result of Bugeaud and Luca [2], it is in principle not possible to determine all triples with the methods currently available.

In contrast to the case of triples we can, in principle, effectively determine all S-Diophantine quadruples for a given set S due to the result of Stewart and Tijdeman [15]. Recently, Szalay and Ziegler [16], established an efficient algorithm to determine all S-Diophantine quadruples for a given set S of primes, provided $|S| = 2$. In particular, the results of Szalay and Ziegler [16], [17], [18], suggest that for $|S| = 2$ no quadruple exists at all.

The aim of this note is to give upper bounds for the number of S-Diophantine quadruples for fixed sets S of r primes. We need the following notations. Let Γ be a multiplicative subgroup of $\mathbb{Q}^*$ of rank r and denote by $A(n,r)$ an upper bound for the number of non-degenerate solutions $(x_1, \dots, x_n) \in \Gamma^n$ to the linear S-unit equation

$$a_1 x_1 + \cdots + a_n x_n = 1\,, \quad a_i \in \mathbb{Q}^*. \tag{1}$$

We call a solution to (1) non-degenerate if no subsum on the left hand side of equation (1) vanishes. With this notation at hand our main result is:

Theorem 1. *Let S be a set of r primes. Then there exist at most*

$$\big(A(5,r) + A(2,r)^2\big)A(3,r)$$

S-Diophantine quadruples. If $r = 2$ or $2 \notin S$, then there exist at most

$$A(5,r)A(3,r)$$

S-Diophantine quadruples.

Using the best estimates for $A(n,r)$ currently available we obtain

Corollary 1. *Let S be a set of r primes. Then there exist at most*

$$\exp(27398 + 5136r)$$

S-Diophantine quadruples.

In the next section we prove Theorem 1 and in the third section we discuss the number of solutions to the S-unit equation (1) and establish Corollary 1.

2 A system of S-unit equations

Assume that (a,b,c,d) is an S-Diophantine quadruple, with $a<b<c<d$. We write,

$$
\begin{aligned}
ab+1&=s_1\,, & ac+1&=s_2\,, & ad+1&=s_3\,,\\
bc+1&=s_4\,, & bd+1&=s_5\,, & cd+1&=s_6\,.
\end{aligned}
$$

With these notations we have

$$
\begin{aligned}
abcd&=s_1s_6-s_1-s_6+1\\
&=s_2s_5-s_2-s_5+1\\
&=s_3s_4-s_3-s_4+1
\end{aligned}
$$

and obtain the following system of S-unit equations

$$
\begin{aligned}
s_1s_6-s_1-s_6-s_2s_5+s_2+s_5&=0\,,\\
s_1s_6-s_1-s_6-s_3s_4+s_3+s_4&=0\,.
\end{aligned}
\tag{2}
$$

Let us consider the first equation more closely and write $y_1=s_1s_6$, $y_2=s_1$, $y_3=s_6$, $y_4=s_2s_5$, $y_5=s_2$ and $y_6=s_5$. Then the first equation of system (2) takes the form

$$y_1-y_2-y_3-y_4+y_5+y_6=0$$

and has at most $A(5,r)$ projective solutions in $\mathbb{P}^5(\Gamma)$ such that no subsum vanishes, where $\Gamma\subset\mathbb{Q}^*$ is the multiplicative group generated by S. Note that each projective solution yields at most one solution (s_1,s_2,s_5,s_6). Indeed, assume (s_1,s_2,s_5,s_6) and (s_1',s_2',s_5',s_6') correspond to the same projective solution. Then there is a rational number $\rho\neq 0$ such that $s_1=\rho s_1'$, $s_6=\rho s_6'$, $s_2=\rho s_2'$, $s_5=\rho s_5'$ and $s_1s_6=\rho s_1's_6'$. But this implies that $s_1s_6=\rho^2s_1's_6'=\rho s_1's_6'$, thus $\rho=1$ and $s_i=s_i'$ for $i=1,2,5,6$.

So we are left to count how many solutions exist with vanishing subsums. Of course there exist no vanishing one-term subsums. Two-term vanishing subsums imply either

- $s_i=s_j$ for $i\neq j$ which is impossible, unless $i,j\in\{3,4\}$ which is excluded, or
- $s_i=s_1s_6>abcd>cd+1\geq s_6\geq s_i$ for some $i\in\{1,2,5,6\}$ which is also a contradiction, or
- $s_i=s_2s_5>abcd>cd+1\geq s_6\geq s_i$ for some $i\in\{1,2,5,6\}$ which is also a contradiction, or
- $s_1s_6=s_2s_5$, which implies $ab+cd+2=s_1+s_6=s_2+s_5=ac+bd+2$; hence, $(c-b)(d-a)=0$; i.e., $d=a$ or $b=c$, again a contradiction.

Therefore no two-term subsums vanish. Since four- and five-term vanishing subsums imply the existence of two- and one-term vanishing subsums, respectively, we are left with the case of three-term vanishing subsums.

Without loss of generality we may assume that the vanishing three-term subsum contains s_1s_6. Thus we distinguish whether s_2s_5 is contained in the vanishing subsum or not. Let us consider the case that s_2s_5 is not contained. Then we have an equation of the from $s_1s_6 = \pm s_i \pm s_j$. Since $s_1 = ab+1 > 2\cdot 1+1 > 2$ we have $s_1s_6 > 2s_6 > s_i + s_j$ and this case yields no solution.

Therefore both s_1s_6 and s_2s_5 are contained in the same vanishing three-term subsum and we are left with four systems of S-unit equations namely

$$\begin{aligned} s_1s_6 - s_5s_2 &= s_1 \quad \text{and} \quad s_6 = s_5 + s_2\,, \\ s_1s_6 - s_5s_2 &= s_6 \quad \text{and} \quad s_1 = s_5 + s_2\,, \\ s_1s_6 - s_5s_2 &= -s_2 \quad \text{and} \quad s_1 + s_6 = s_5\,, \\ s_1s_6 - s_5s_2 &= -s_5 \quad \text{and} \quad s_1 + s_6 = -s_2\,. \end{aligned} \tag{3}$$

Note that only the first equation of (3) is possible since by assumption $s_1 < s_2 < s_5 < s_6$. Let $y_1 = s_1s_6$, $y_2 = s_5s_2$ and $y_3 = s_1$. Then the S-unit equation

$$y_1 - y_2 = y_3$$

has at most $A(2,r)$ projective solutions $(y_1, y_2, y_3) \in \mathbb{P}^2(\Gamma)$. Note that all solutions that yield S-Diophantine quadruples are non-degenerate, since a vanishing subsum would imply either $s_1s_6 = 0$ or $s_2s_5 = 0$ or $s_1 = 0$. Each projective solution yields only one possibility for s_6. Indeed, assume that (s_1, s_2, s_5, s_6) and (s_1', s_2', s_5', s_6') yield the same projective solution. Then there exists $\rho \in \mathbb{Q}^*$ such that $s_1s_6 = \rho s_1's_6' = s_1s_6'$, since $s_1 = \rho s_1'$, i.e. $s_6 = s_6'$. We have now at most $A(2,r)$ possible values for s_6; i.e., we are reduced to at most $A(2,r)$ equations of the form

$$a = s_5 + s_2$$

with $a = s_6 \neq 0$ fixed. Thus, system (3) yields at most $A(2,r)^2$ solutions.

In view of the second statement of Theorem 1 we note that any equation of system (3) cannot have a solution if $2 \notin S$. Otherwise s_6 is odd but $s_5 + s_2$ would be even. In case of $r = 2$, this implies $S = \{2, p\}$ and the equation $s_6 = s_5 + s_2$ turns into

$$2^{\alpha_6}p^{\beta_6} = 2^{\alpha_5}p^{\beta_5} + 2^{\alpha_2}p^{\beta_2}. \tag{4}$$

Considering 2-adic and p-adic valuations, equation (4) reduces to the Diophantine equation

$$2^x - p^y = \pm 1\,.$$

By Mihăilescu's solution of Catalan's equation [11], only $p = 3$ is possible. On the other hand, Szalay and Ziegler [16] showed that no $\{2,3\}$-Diophantine quadruple exists.

Altogether, we have proved the following result.

Lemma 1. *The first S-unit equation in* (2) *has at most $A(5,r) + A(2,r)^2$ solutions. If $r = 2$ or $2 \notin S$, then there exist at most $A(5,r)$ solutions.*

Now, we turn to the second equation of system (2). By Lemma 1, the first equation in (2) yields at most $A(5,r)+A(2,r)^2$ or $A(5,r)$ many possibilities for the pair (s_1,s_6) respectively. Thus, we assume that the second equation of system (2) is of the form

$$s_3s_4 - s_3 - s_4 = a \quad \text{with } a \in \mathbb{Q} \text{ fixed.} \tag{5}$$

But S-unit equation (5) has at most $A(3,r)$ solutions provided $a \neq 0$. Indeed no degenerate solution exists since a vanishing subsum on the left side of equation (5) would imply either

- $s_3s_4 = s_3$ and therefore $s_4 = 1$, or
- $s_3s_4 = s_4$ and therefore $s_3 = 1$, or
- $s_3 + s_4 = 0$ and therefore $s_3s_4 < 0$.

Let us note that $a = s_6s_1 - s_6 - s_1 > 2s_6 - s_6 - s_1 > 0$, and therefore we have proved the following lemma.

Lemma 2. *The Diophantine system* (2) *has at most* $(A(5,r) + A(2,r)^2)A(3,r)$ *solutions. If* $r = 2$ *or* $2 \notin S$*, then there exist at most* $A(5,r)A(3,r)$ *solutions.*

In order to prove Theorem 1 it remains to prove that for fixed integers $s_1,\dots,s_6$ there exists at most one quadruple (a,b,c,d). Since

$$a = \sqrt{\frac{(s_1-1)(s_2-1)}{s_4-1}}, \qquad b = \sqrt{\frac{(s_1-1)(s_4-1)}{s_2-1}},$$

$$c = \sqrt{\frac{(s_2-1)(s_4-1)}{s_1-1}}, \qquad d = \sqrt{\frac{(s_5-1)(s_6-1)}{s_4-1}},$$

the proof of Theorem 1 is complete.

3 Proof of Corollary 1

A look through the vast literature on S-unit equations shows that for S-unit equations over the rationals the best result is due to Evertse [5] provided $|S| = 2$ and due to Amoroso and Viada [1] in the general case. Therefore we may assume $A(2,r) = 3 \cdot 7^{3+2r}$ and $A(n,r) = (8n)^{4n^4(n+r+1)}$. A look at the proof of the bound for $A(n,r)$ in [1] shows that this bound is derived by the recursive relation

$$A(n,r) \leq 2^n A(n-1,r)B(n,r+1),$$

where $B(n,r) = (8n)^{6n^3(n+r)}$. Note that this recursive estimate already appears in [7]. However, recursively computing $A(n,r)$ we obtain

$$A(3,r) \leq 8 \cdot 3 \cdot 7^{3+2r} \cdot 24^{162(4+r)} < \exp(2069 + 518.8r).$$

Continuing these computations we arrive at

$$A(5,r) < \exp(25329 + 4616.3r).$$

With these numbers plugged into Theorem 1, we obtain Corollary 1.

Remark 1. Let us note that directly applying the bounds due to Evertse [5] and Amoroso and Viada [1] would yield the slightly worse bound $\exp(73801+15378r)$ for the number of S-Diophantine quadruples. A closer inspection of the computation of the quantity $B(n,r)$ due to Amoroso and Viada [1] and Evertse et al. [7] would further improve the bounds also in view of the new improvements of the Subspace Theorem due to Evertse and Feretti [6]. But we are afraid that the gain is too small for such an effort.

Acknowledgement

We thank the anonymous referee for his valuable suggestions on treating the three--term vanishing subsums in the proof of Theorem 1.

The first author worked on this paper in Fall of 2013 as a long term guest of the Special Semester on Applications of Algebra and Number Theory at the RICAM, Linz, Austria. He thanks Arne Winterhof for the invitation to participate in this program and RICAM for hospitality. The first author was also supported in part by Projects PAPIIT IN104512, CONACyT Mexico-France 193539, CONACyT Mexico-India 163787, and a Marcos Moshinsky Fellowship. The second author was supported by the Austrian Science Fund (FWF) under the project P 24801-N26.

References

[1] F. Amoroso, E. Viada: Small points on subvarieties of a torus. *Duke Math. J.* 150 (3) (2009) 407–442.

[2] Y. Bugeaud, F. Luca: A quantitative lower bound for the greatest prime factor of $(ab+1)(bc+1)(ca+1)$. *Acta Arith.* 114 (3) (2004) 275–294.

[3] P. Corvaja, U. Zannier: On the greatest prime factor of $(ab+1)(ac+1)$. *Proc. Amer. Math. Soc.* 131 (6) (2003) 1705–1709. (electronic)

[4] P. Erdős, P. Turan: On a Problem in the Elementary Theory of Numbers. *Amer. Math. Monthly* 41 (10) (1934) 608–611.

[5] J.-H. Evertse: On equations in S-units, the Thue-Mahler equation. *Invent. Math.* 75 (3) (1984) 561–584.

[6] J.-H. Evertse, R. G. Ferretti: A further improvement of the quantitative subspace theorem. *Ann. of Math. (2)* 177 (2) (2013) 513–590.

[7] J.-H. Evertse, H. P. Schlickewei, W. M. Schmidt: Linear equations in variables which lie in a multiplicative group. *Ann. of Math. (2)* 155 (3) (2002) 807–836.

[8] K. Győry, A. Sárközy, C. L. Stewart: On the number of prime factors of integers of the form $ab+1$. *Acta Arith.* 74 (4) (1996) 365–385.

[9] S. Hernández, F. Luca: On the largest prime factor of $(ab+1)(ac+1)(bc+1)$. *Bol. Soc. Mat. Mexicana (3)* 9 (2) (2003) 235–244.

[10] F. Luca: On the greatest common divisor of $u-1$ and $v-1$ with u and v near S-units. *Monatsh. Math.* 146 (3) (2005) 239–256.

[11] P. Mihăilescu: Primary cyclotomic units and a proof of Catalan's conjecture. *J. Reine Angew. Math.* 572 (2004) 167–195.

[12] A. Sárközy, C. L. Stewart: On divisors of sums of integers. II. *J. Reine Angew. Math.* 365 (1986) 171–191.

[13] A. Sárközy, C. L. Stewart: On divisors of sums of integers. V. *Pacific J. Math.* 166 (2) (1994) 373–384.

[14] A. Sárközy, C. L. Stewart: On prime factors of integers of the form $ab+1$. *Publ. Math. Debrecen* 56 (3–4) (2000) 559–573. Dedicated to Professor Kálmán Győry on the occasion of his 60th birthday.

[15] C. L. Stewart, R. Tijdeman: On the greatest prime factor of $(ab+1)(ac+1)(bc+1)$. *Acta Arith.* 79 (1) (1997) 93–101.

[16] L. Szalay, V. Ziegler: S-diophantine quadruples with $S=\{2,q\}$. (in preperation)

[17] L. Szalay, V. Ziegler: On an S-unit variant of Diophantine m-tuples. *Publ. Math. Debrecen* 83 (1–2) (2013) 97–121.

[18] L. Szalay, V. Ziegler: S-diophantine quadruples with two primes congruent 3 modulo 4. *Integers* 13 (2013). Paper No. A80, 9pp.

Authors' addresses:
F. Luca: Mathematical Institute, UNAM Juriquilla, Juriquilla, 76230 Santiago de Querétaro, Querétaro de Arteaga, México, and
School of Mathematics, University of the Witwatersrand, P. O. Box Wits 2050, South Africa

E-mail: fluca@matmor.unam.mx

V. Ziegler: Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Altenbergerstr. 69, A-4040 Linz, Austria

E-mail: volker.ziegler@ricam.oeaw.ac.at

Received: 3 February, 2014
Accepted for publication: 24 March, 2014
Communicated by: Attila Bérczes

Communications in Mathematics 22 (2014) 57–69

Existence of entropy solutions for degenerate quasilinear elliptic equations in L^1

Albo Carlos Cavalheiro

Abstract. In this article, we prove the existence of entropy solutions for the Dirichlet problem

$$(P)\begin{cases} -\operatorname{div}[\omega(x)\mathcal{A}(x,u,\nabla u)] = f(x) - \operatorname{div}(G), & \text{in } \Omega \\ u(x) = 0, & \text{on } \partial\Omega \end{cases}$$

where Ω is a bounded open set of $\mathbb{R}^N$, $N \geq 2$, $f \in L^1(\Omega)$ and $G/\omega \in [L^{p'}(\Omega,\omega)]^N$.

1 Introduction

The main purpose of this article (see Theorem 2) is to establish the existence of entropy solutions for the Dirichlet problem

$$(P)\begin{cases} -\operatorname{div}[\omega(x)\mathcal{A}(x,u,\nabla u)] = f(x) - \operatorname{div}(G) & \text{in } \Omega \\ u(x) = 0 & \text{on } \partial\Omega \end{cases}$$

where $\Omega \subset \mathbb{R}^N$ is a bounded open set, ω is a weight function (i.e., a locally integrable function on $\mathbb{R}^N$ such that $0 < \omega(x) < \infty$ a.e. $x \in \mathbb{R}^N$), $f \in L^1(\Omega)$, $G = (g_1, \dots, g_N)$ with $G/\omega \in [L^{p'}(\Omega,\omega)]^N$, and the function

$$\mathcal{A}\colon \Omega \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}^N$$

satisfies the following conditions:

(H1) $x \mapsto \mathcal{A}(x,s,\xi)$ is measurable on Ω for all $(s,\xi) \in \mathbb{R} \times \mathbb{R}^N$ and $(s,\xi) \mapsto \mathcal{A}(x,s,\xi)$ is continuous on $\mathbb{R} \times \mathbb{R}^N$ for almost all $x \in \Omega$.

(H2) $\langle \mathcal{A}(x,s,\xi_1) - \mathcal{A}(x,s,\xi_2), \xi_1 - \xi_2 \rangle > 0$, whenever $\xi_1, \xi_2 \in \mathbb{R}^N$, $\xi_1 \neq \xi_2$ (where $\langle \cdot,\cdot \rangle$ denotes the usual inner product in $\mathbb{R}^N$).

2010 MSC: 35J70, 35J60
Key words: Degenerate elliptic equations, entropy solutions, weighted Sobolev spaces

(H3) $\langle \mathcal{A}(x,s,\xi),\xi\rangle \geq \lambda\,|\xi|^p$, with $1<p<\infty$, and $\lambda>0$.

(H4) $|\mathcal{A}(x,s,\xi)| \leq K(x) + h_1(x)\,|s|^{p/p'} + h_2(x)\,|\xi|^{p/p'}$, where K, h_1 and h_2 are positive functions, with $h_1 \in L^\infty(\Omega)$, $h_2 \in L^\infty(\Omega)$ and $K \in L^{p'}(\Omega,\omega)$ (where $1/p + 1/p' = 1$).

If $f/\omega \in L^{p'}(\Omega,\omega)$ (with $1<p<\infty$), the problem (P) has been studied in [4], and in this case the problem (P) has a solution $u \in W^{1,p}(\Omega,\omega)$. However, since $L^1(\Omega)$ is not a subspace of $W^{-1,p'}(\Omega,\omega)$ so when we want to consider $f \in L^1(\Omega)$ a different theory is needed.

In [1], a new concept of solution has been introduced for the elliptic equation

$$\begin{cases} -\operatorname{div}[a(x,\nabla u)] = f(x) & \text{in } \Omega,\\ u = 0 & \text{on } \partial\Omega \end{cases}$$

(when $f \in L^1(\Omega)$) namely entropy solution. In [3] the author studied the degenerate elliptic equation $Lu = f$, where L is a degenerate elliptic operator in divergence form, i.e.,

$$Lu = -\sum_{i,j=1}^{n} D_j(a_{ij}(x)D_i u),$$

and $f \in L^1(\Omega)$. Note that, in the proof of our main result, many ideas have been adapted from [1] and [3].

For degenerate partial differential equations, i.e., equations with various types of singularities in the coefficients, it is natural to look for solutions in weighted Sobolev spaces (see [5], [6], [8], [9] and [13]).

A class of weights, which is particularly well understood, is the class of A_p weights that was introduced by B. Muckenhoupt in the early 1970s (see [11]).

We propose to solve the problem (P) by approximation with variational solutions: we take $f_n \in C_0^\infty(\Omega)$ such that $f_n \to f$ in $L^1(\Omega)$, $G_n/\omega \in [L^{p'}(\Omega,\omega)]^N$ such that $G_n/\omega \to G/\omega$ in $[L^{p'}(\Omega,\omega)]^N$, we find a solution $u_n \in W_0^{1,p}(\Omega,\omega)$ for the problem with right-hand side f_n and G_n, and we will try to pass to the limit as $n\to\infty$.

2 Definitions and basic results

By a weight we shall mean a locally integrable function ω on $\mathbb{R}^N$ such that $0<\omega(x)<\infty$ for a.e. $x\in\mathbb{R}^N$. Every weight ω gives rise to a measure on the measurable subsets of $\mathbb{R}^N$ through integration. This measure will be denoted by μ. Thus, $\mu(E) = \displaystyle\int_E \omega(x)\,\mathrm{d}x$ for measurable sets $E\subset\mathbb{R}^N$.

Definition 1. Let $1\leq p<\infty$. A weight ω is said to be an A_p-weight, if there is a positive constant $C = C(p,\omega)$ such that for every ball $B\subset\mathbb{R}^N$

$$\left(\frac{1}{|B|}\int_B \omega(x)\,\mathrm{d}x\right)\left(\frac{1}{|B|}\int_B \omega^{1/(1-p)}(x)\,\mathrm{d}x\right)^{p-1} \leq C \quad \text{if } p>1,$$

$$\left(\frac{1}{|B|}\int_B \omega(x)\,\mathrm{d}x\right)\left(\operatorname*{ess\,sup}_{x\in B}\frac{1}{\omega(x)}\right) \leq C \quad \text{if } p=1,$$

where $|\cdot|$ denotes the N-dimensional Lebesgue measure in $\mathbb{R}^N$.

If $1 < q \leq p$, then $A_q \subset A_p$ (see [8], [9] or [14] for more information about A_p-weights). As an example of an A_p-weight, the function $\omega(x) = |x|^\alpha$, $x \in \mathbb{R}^N$, is in A_p if and only if, $-N < \alpha < N(p-1)$ (see [12], Chapter IX, Corollary 4.4). If $\varphi \in BMO(\mathbb{R}^N)$ then $\omega(x) = \mathrm{e}^{\alpha\,\varphi(x)} \in A_2$ for some $\alpha > 0$ (see [12]).

Remark 1. If $\omega \in A_p$, $1 < p < \infty$, then

$$\left(\frac{|E|}{|B|}\right)^p \leq C\,\frac{\mu(E)}{\mu(B)}$$

for all measurable subsets E of B (see *15.5 strong doubling property* in [9]). Therefore if $\mu(E) = 0$ then $|E| = 0$. Thus, if $\{u_n\}$ is a sequence of functions defined in B and $u_n \to u$ μ-a.e. then $u_n \to u$ a.e..

Definition 2. Let ω be a weight. We shall denote by $L^p(\Omega,\omega)$ $(1 \leq p < \infty)$ the Banach space of all measurable functions f defined in Ω for which

$$\|f\|_{L^p(\Omega,\omega)} = \left(\int_\Omega |f(x)|^p \omega(x)\,\mathrm{d}x\right)^{1/p} < \infty\,.$$

We denote $[L^{p'}(\Omega,\omega)]^N = L^{p'}(\Omega,\omega) \times \cdots \times L^{p'}(\Omega,\omega)$.

Remark 2. If $\omega \in A_p$, $1 < p < \infty$, then since $\omega^{-1/(p-1)}$ is locally integrable, we have $L^p(\Omega,\omega) \subset L^1_{\mathrm{loc}}(\Omega)$ (see [14], Remark 1.2.4). It thus makes sense to talk about weak derivatives of functions in $L^p(\Omega,\omega)$.

Definition 3. Let $\Omega \subset \mathbb{R}^N$ a bounded open set, $1 < p < \infty$, k a nonnegative integer and $\omega \in A_p$. We shall denote by $W^{k,p}(\Omega,\omega)$, the weighted Sobolev spaces, the set of all functions $u \in L^p(\Omega,\omega)$ with weak derivatives $D^\alpha u \in L^p(\Omega,\omega)$, $1 \leq |\alpha| \leq k$. The norm in the space $W^{k,p}(\Omega,\omega)$ is defined by

$$\|u\|_{W^{k,p}(\Omega,\omega)} = \left(\int_\Omega |u(x)|^p \omega(x)\,\mathrm{d}x + \sum_{1\leq|\alpha|\leq k} \int_\Omega |D^\alpha u(x)|^p \omega(x)\,\mathrm{d}x\right)^{1/p}. \tag{1}$$

We also define the space $W^{k,p}_0(\Omega,\omega)$ as the closure of $C^\infty_0(\Omega)$ with respect to the norm

$$\|u\|_{W^{k,p}_0(\Omega,\omega)} = \left(\sum_{1\leq|\alpha|\leq k} \int_\Omega |D^\alpha u(x)|^p \omega(x)\,\mathrm{d}x\right)^{1/p}.$$

The dual space of $W^{1,p}_0(\Omega,\omega)$ is the space $[W^{1,p}_0(\Omega,\omega)]^* = W^{-1,p'}(\Omega,\omega)$,

$$W^{-1,p'}(\Omega,\omega) = \left\{T = f - \operatorname{div}(G) : G = (g_1,\ldots,g_N), \frac{f}{\omega}, \frac{g_j}{\omega} \in L^{p'}(\Omega,\omega)\right\}.$$

It is evident that a weight function ω which satisfies $0 < C_1 \leq \omega(x) \leq C_2$, for a.e. $x \in \Omega$, gives nothing new (the space $\mathrm{W}^{k,p}(\Omega,\omega)$ is then identical with the classical Sobolev space $\mathrm{W}^{k,p}(\Omega)$). Consequently, we study all such weight function ω that either vanish in $\Omega \cup \partial\Omega$ or increase to infinity (or both).

We need the following basic result.

Theorem 1. *(The weighted Sobolev inequality) Let $\Omega \subset \mathbb{R}^N$ be a bounded open set and let ω be an A_p-weight, $1 < p < \infty$. Then there exist positive constants C_Ω and δ such that for all $f \in C_0^\infty(\Omega)$ and $1 \le \eta \le N/(N-1) + \delta$*

$$\|f\|_{L^{\eta p}(\Omega,\omega)} \le C_\Omega \|\nabla f\|_{L^p(\Omega,\omega)}. \tag{2}$$

Proof. See [6], Theorem 1.3. □

Definition 4. We say that $u \in \mathcal{T}_0^{1,p}(\Omega,\omega)$ if $T_k(u) \in W_0^{1,p}(\Omega,\omega)$, for all $k > 0$, where the function $T_k \colon \mathbb{R} \to \mathbb{R}$ is defined by

$$T_k(s) = \begin{cases} s, & \text{if } |s| \le k \\ k \operatorname{sign}(s), & \text{if } |s| > k. \end{cases}$$

Remark 3. (i) Note that for given $h > 0$ and $k > 0$ we have

$$T_h(u - T_k(u)) = \begin{cases} 0, & \text{if } |u| \le k \\ (|u| - k) \operatorname{sign}(u), & \text{if } k < |u| \le k + h \\ h \operatorname{sign}(u), & \text{if } |u| > k + h. \end{cases}$$

And if $\alpha \in \mathbb{R}$, $\alpha \ne 0$, we have $T_k(\alpha u) = \alpha\, T_{k/|\alpha|}(u)$.

(ii) If $u \in W^{1,1}_{\mathrm{loc}}(\Omega,\omega)$ then we have

$$\nabla T_k(u) = \chi_{\{|u|<k\}} \nabla u$$

where χ_E denotes the characteristic function of a measurable set $E \subset \mathbb{R}^N$.

Definition 5. Let $f \in L^1(\Omega)$, $G/\omega \in [L^{p'}(\Omega,\omega)]^N$ and $u \in \mathcal{T}_0^{1,p}(\Omega,\omega)$. We say that u is an entropy solution to problem (P) if

$$\int_\Omega \big\langle \mathcal{A}(x,u,\nabla u), \nabla T_k(u-\varphi)\big\rangle\, \omega \,\mathrm{d}x = \int_\Omega f T_k(u-\varphi)\,\mathrm{d}x + \int_\Omega \big\langle G, \nabla T_k(u-\varphi)\big\rangle\,\mathrm{d}x \tag{3}$$

for all $k > 0$ and all $\varphi \in W_0^{1,p}(\Omega,\omega) \cap L^\infty(\Omega)$.

We recall that the gradient of u which appears in (3) is defined as in Remark 2.8 of [3], that is to say that $\nabla u = \nabla T_k(u)$ on the set where $|u| < k$.

Remark 4. Note that if $u_1, u_2 \in W_0^{1,p}(\Omega,\omega)$ then

$$\varphi = T_k(u_1 + u_2) \in W_0^{1,p}(\Omega,\omega) \cap L^\infty(\Omega)$$

and we have

$$\nabla\varphi = \nabla T_k(u_1+u_2) = \nabla(u_1+u_2)\,\chi_{\{|u_1+u_2|\le k\}}.$$

Definition 6. Let $0 < p < \infty$ and let ω be a weight function. We define the weighted Marcinkiewicz space $\mathcal{M}^p(\Omega,\omega)$ as the set of all measurable functions $f \colon \Omega \to \mathbb{R}$ such that the function

$$\Gamma_f(k) = \mu(\{x \in \Omega : |f(x)| > k\}),\ k > 0,$$

satisfies an estimate of the form $\Gamma_f(k) \le Ck^{-p}$, $0 < C < \infty$.

Remark 5. If $1 \leq q < p$ and $\Omega \subset \mathbb{R}^N$ is a bounded set, we have that

$$L^p(\Omega, \omega) \subset \mathcal{M}^p(\Omega, \omega) \text{ and } \mathcal{M}^p(\Omega, \omega) \subset L^q(\Omega, \omega)$$

(the proof follows the lines of Theorem 2.18.8 in [10]).

Lemma 1. *Let $u \in \mathcal{T}_0^{1,p}(\Omega, \omega)$ and $\omega \in A_p$, $1 < p < \infty$, be such that*

$$\frac{1}{k} \int_{\{|u|<k\}} |\nabla u|^p \omega \, \mathrm{d}x \leq M \,, \tag{4}$$

for every $k > 0$. Then $u \in \mathcal{M}^{p_1}(\Omega, \omega)$, where $p_1 = \eta\,(p-1)$ (where η is the constant in Theorem 1). More precisely, there exists $C > 0$ such that $\Gamma_u(k) \leq C M^\eta k^{-p_1}$.

Proof. See Lemma 3.3 in [3]. □

Lemma 2. *Let $u \in \mathcal{T}_0^{1,p}(\Omega, \omega)$, where $\omega \in A_p$, $1 < p < \infty$, be such that*

$$\frac{1}{k} \int_{\{|u|<k\}} |\nabla u|^p \omega \, \mathrm{d}x \leq M \,,$$

for every $k > 0$. Then $|\nabla u| \in \mathcal{M}^{p_2}(\Omega, \omega)$, where $p_2 = p\,p_1/(p_1+1)$ (with η as in Lemma 2 and $p_1 = \eta(p-1)$). More precisely, there exists $C > 0$ such that

$$\Gamma_k(|\nabla u|) \leq C M^{(p_1+\eta)/(p_1+1)} k^{-p_2}.$$

Proof. See Lemma 3.4 in [3]. □

3 Main Result

In this section, we prove the main result of this paper. We need the following result.

Lemma 3. *Let $\omega \in A_p$, $1 < p < \infty$ and a sequence $\{u_n\}$, $u_n \in W_0^{1,p}(\Omega, \omega)$ satisfies*

(i) $u_n \rightharpoonup u$ in $W_0^{1,p}(\Omega, \omega)$ and μ-a.e. in Ω;

(ii) $\displaystyle\int_\Omega \langle \mathcal{A}(x, u_n, \nabla u_n) - \mathcal{A}(x, u_n, \nabla u), \nabla(u_n - u)\rangle \, \omega \, \mathrm{d}x \to 0$ with $n \to \infty$.

Then $u_n \to u$ in $W_0^{1,p}(\Omega, \omega)$.

Proof. The proof of this lemma follows the lines of Lemma 5 in [2]. □

Theorem 2. *Let $\omega \in A_p$, $1 < p < \infty$, and $\mathcal{A}(x, s, \xi)$ satisfies the conditions (H1), (H2), (H3) and (H4). Then, there exists an entropy solution u of problem (P). Moreover, $u \in \mathcal{M}^{p_1}(\Omega, \omega)$ and $|\nabla u| \in \mathcal{M}^{p_2}(\Omega, \omega)$, with $p_1 = \eta\,(p-1)$ and $p_2 = p_1\,p/(p_1+1)$ (where η is the constant in Theorem 1).*

Proof. Considering a sequence $\{f_n\}$, $f_n \in C_0^\infty(\Omega)$, which

$$f_n \to f \text{ in } L^1(\Omega) \text{ and } \|f_n\|_{L^1(\Omega)} \leq \|f\|_{L^1(\Omega)}\,,$$

and a sequence $\{G_n\}$, with $G_n/\omega \in [L^{p'}(\Omega,\omega)]^N$ such that

$$\frac{G_n}{\omega} \to \frac{G}{\omega} \text{ in } [L^{p'}(\Omega,\omega)]^N \text{ and } \left\|\frac{|G_n|}{\omega}\right\|_{L^{p'}(\Omega,\omega)} \leq \left\|\frac{|G|}{\omega}\right\|_{L^{p'}(\Omega,\omega)}.$$

For each n, there exists a solution $u_n \in W_0^{1,p}(\Omega,\omega)$ of the Dirichlet problem

$$(P_n)\begin{cases} -\operatorname{div}[\omega(x)\mathcal{A}(x,u_n,\nabla u_n)] = f_n(x) - \operatorname{div}(G_n), & \text{in } \Omega \\ u_n(x) = 0, & \text{on } \partial\Omega \end{cases}$$

(by Theorem 1.1 in [4]) that is,

$$\int_\Omega \omega\, \langle \mathcal{A}(x,u_n,\nabla u_n), \nabla\varphi\rangle \,\mathrm{d}x = \int_\Omega f_n\varphi\,\mathrm{d}x + \int_\Omega \langle G_n,\nabla\varphi\rangle\,\mathrm{d}x\,, \tag{5}$$

for all $\varphi \in W_0^{1,p}(\Omega,\omega)$. For $\varphi = T_k(u_n)$ we obtain in (5) that

$$\int_\Omega \omega\, \langle \mathcal{A}(x,u_n,\nabla u_n), \nabla T_k(u_n)\rangle \,\mathrm{d}x = \int_\Omega f_n T_k(u_n)\,\mathrm{d}x + \int_\Omega \langle G_n,\nabla T_k(u_n)\rangle\,\mathrm{d}x. \tag{6}$$

Using (H3) and Remark 3 (ii) we have,

$$\begin{aligned}\int_\Omega \omega\, \langle \mathcal{A}(x,u_n,\nabla u_n), \nabla T_k(u_n)\rangle \,\mathrm{d}x &= \int_\Omega \omega\, \langle \mathcal{A}(x,u_n,\nabla T_k(u_n)), \nabla T_k(u_n)\rangle \,\mathrm{d}x \\ &\geq \lambda \int_\Omega |\nabla T_k(u_n)|^p\omega\,\mathrm{d}x.\end{aligned}$$

We also have

$$\left|\int_\Omega f_n T_k(u_n)\,\mathrm{d}x\right| \leq \int_\Omega |f_n|\,|T_k(u_n)|\,\mathrm{d}x \leq k\|f_n\|_{L^1(\Omega)} \leq k\|f\|_{L^1(\Omega)}\,,$$

and using Young's inequality there exists a constant $C_1 > 0$ such that

$$\begin{aligned}\left|\int_\Omega \langle G_n,\nabla T_k(u_n)\rangle\,\mathrm{d}x\right| &\leq \int_\Omega \left|\frac{G_n}{\omega}\right| |\nabla T_k(u_n)|\,\omega\,\mathrm{d}x \\ &\leq \left(\int_\Omega \left|\frac{G_n}{\omega}\right|^{p'}\omega\,\mathrm{d}x\right)^{1/p'}\left(\int_\Omega |\nabla T_k(u_n)|^p\omega\,\mathrm{d}x\right)^{1/p} \\ &\leq \frac{\lambda}{2}\int_\Omega |\nabla T_k(u_n)|^p\omega\,\mathrm{d}x + C_1\int_\Omega \left|\frac{G_n}{\omega}\right|^{p'}\omega\,\mathrm{d}x \\ &\leq \frac{\lambda}{2}\int_\Omega |\nabla T_k(u_n)|^p\omega\,\mathrm{d}x + C_1\int_\Omega \left|\frac{G}{\omega}\right|^{p'}\omega\,\mathrm{d}x\,.\end{aligned}$$

Hence in (6) we obtain

$$\lambda \int_\Omega |\nabla T_k(u_n)|^p \omega \,dx \le k\,\|f\|_{L^1(\Omega)} + \frac{\lambda}{2}\int_\Omega |\nabla T_k(u_n)|^p \omega \,dx + C_1 \int_\Omega \left|\frac{G}{\omega}\right|^{p'} \omega \,dx\,,$$

and

$$\begin{aligned}\int_\Omega |\nabla T_k(u_n)|^p \omega \,dx &\le \frac{k}{\lambda}\bigg(\|f\|_{L^1(\Omega)} + C_1 \left\|\frac{G}{\omega}\right\|^{p'}_{L^{p'}(\Omega,\omega)}\bigg) \\ &= C_2\, k\,, \quad \text{for all } k > 0\,. \end{aligned} \tag{7}$$

By Lemma 1 and Lemma 2, we have that the sequence $\{u_n\}$ is bounded in $\mathcal{M}^{p_1}(\Omega,\omega)$ (with $p_1 = \eta\,(p-1)$ and $\{|\nabla u_n|\}$ is bounded in $\mathcal{M}^{p_2}(\Omega,\omega)$ (with $p_2 = p_1\,p/(p_1+1)$). Moreover, $\{u_n\}$ is a Cauchy sequence in μ-measure. Consequently, there exists a function u and a subsequence, that we will still denote by $\{u_n\}$, such that

$$u_n \to u \quad \mu\text{-a.e. in } \Omega, \tag{8}$$

and $u_n \to u$ a.e. in Ω (by Remark 1).

Using (7) and (8), we have

$$\begin{aligned} &T_k(u_n) \rightharpoonup T_k(u) \quad \text{weakly in } W_0^{1,p}(\Omega,\omega), \\ &T_k(u_n) \to T_k(u) \quad \text{strongly in } L^p(\Omega,\omega) \text{ and } \mu\text{-a.e. in } \Omega, \end{aligned} \tag{9}$$

for all $k > 0$. Hence $T_k(u) \in W_0^{1,p}(\Omega,\omega)$.

Furthermore, by the weak lower semicontinuity of the norm $W_0^{1,p}(\Omega,\omega)$, we have that (7) still holds for u, that is,

$$\int_\Omega |\nabla T_k(u)|^p \omega \,dx \le C_2\,k\,.$$

Applying Lemma 1 and Lemma 2, we have that $u \in \mathcal{M}^{p_1}(\Omega,\omega)$ (with $p_1 = \eta(p-1)$) and $|\nabla u| \in \mathcal{M}^{p_2}(\Omega,\omega)$ (with $p_2 = p_1\,p/(p_1+1)$).

- We need to show that $T_k(u_n) \to T_k(u)$ strongly in $W_0^{1,p}(\Omega,\omega)$, for all $k > 0$.

Let $h > k$ and applying (5) with function

$$\varphi_n = T_{2k}(u_n - T_h(u_n) + T_k(u_n) - T_k(u))\,,$$

we get

$$\int_\Omega \omega\,\langle \mathcal{A}(x,u_n,\nabla u_n), \nabla\varphi_n\rangle \,dx = \int_\Omega f_n \varphi_n \,dx + \int_\Omega \langle G_n, \nabla\varphi_n\rangle \,dx\,. \tag{10}$$

If we set $M = 4k+h$, we have $\nabla\varphi_n = 0$ for $|u_n| > M$. Hence, since condition (H3) implies that $\mathcal{A}(x,s,0) = 0$, we can write

$$\int_\Omega \omega\,\langle \mathcal{A}(x,T_M(u_n),\nabla T_M(u_n)), \nabla\varphi_n\rangle \,dx = \int_\Omega f_n \varphi_n \,dx + \int_\Omega \langle G_n, \nabla\varphi_n\rangle \,dx\,. \tag{11}$$

In the left-hand side of (11), we have

$$\begin{aligned}&\int_\Omega \omega\big\langle \mathcal{A}(x,T_M(u_n),\nabla T_M(u_n)),\nabla T_{2k}(u_n-T_h(u_n)+T_k(u_n)-T_k(u))\big\rangle\,\mathrm{d}x\\&=\int_{\{|u_n|\le k\}}\omega\big\langle \mathcal{A}(x,T_M(u_n),\nabla T_M(u_n)),\nabla T_{2k}(u_n-T_h(u_n)+T_k(u_n)-T_k(u))\big\rangle\,\mathrm{d}x\\&+\int_{\{|u_n|> k\}}\omega\big\langle \mathcal{A}(x,T_M(u_n),\nabla T_M(u_n)),\nabla T_{2k}(u_n-T_h(u_n)+T_k(u_n)-T_k(u))\big\rangle\,\mathrm{d}x\end{aligned}\tag{12}$$

(a) If $|u_n|\le k$. Since $h>k$, if $|u_n|\le k<h$, then $T_h(u_n)=T_k(u_n)=u_n$. Hence,

$$u_n-T_h(u_n)+T_k(u_n)-T_k(u)=u_n-T_k(u)\,.$$

We also have that $|u_n-u|\le 2k$. Then, since $\nabla T_M(u_n)=\nabla T_k(u_n)$ (because $|u_n|\le k<M$),

$$\begin{aligned}&\int_{\{|u_n|\le k\}}\omega\big\langle \mathcal{A}(x,T_M(u_n),\nabla T_M(u_n)),\nabla T_{2k}(u_n-T_h(u_n)+T_k(u_n)-T_k(u))\big\rangle\,\mathrm{d}x\\&\qquad=\int_{\{|u_n|\le k\}}\omega\big\langle \mathcal{A}(x,T_k(u_n),\nabla T_k(u_n)),\nabla (T_k(u_n)-T_k(u))\big\rangle\,\mathrm{d}x\\&\qquad=\int_\Omega\omega\big\langle \mathcal{A}(x,T_k(u_n),\nabla T_k(u_n)),\nabla (T_k(u_n)-T_k(u))\big\rangle\,\mathrm{d}x\,.\end{aligned}$$

(b) If $|u_n|>k$. Since u_n, $T_k(u_n)$ and $T_k(u)$ are in $W_0^{1,p}(\Omega,\omega)$, if

$$|u_n-T_h(u_n)+T_k(u_n)-T_k(u)|\le 2k\,,$$

we obtain

$$\begin{aligned}\nabla T_{2k}(u_n-T_h(u_n)+T_k(u_n)-T_k(u))&=\nabla(u_n-T_h(u_n)+T_k(u_n)-T_k(u))\\&=\nabla u_n-\nabla T_h(u_n)+\nabla T_k(u_n)-\nabla T_k(u)\\&=\nabla u_n-\nabla T_h(u_n)-\nabla T_k(u)\end{aligned}$$

(because $\nabla T_k(u_n)=0$ if $|u_n|>k$). There are two possible cases as follows:

(i) If $k<|u_n|<h$, we have $\nabla T_h(u_n)=\nabla u_n$. Then

$$\nabla T_{2k}(u_n-T_h(u_n)+T_k(u_n)-T_k(u))=-\nabla T_k(u).$$

(ii) If $h<|u_n|\le M$, we have that $\nabla T_h(u_n)=0$. Then

$$\nabla T_{2k}(u_n-T_h(u_n)+T_k(u_n)-T_k(u))=\nabla u_n-\nabla T_k(u)=\nabla T_M(u_n)-\nabla T_k(u).$$

Since $\langle \mathcal{A}(x,s,\xi),\xi\rangle\ge\lambda|\xi|^p\ge 0$, in both cases we obtain

$$\begin{aligned}&\big\langle \mathcal{A}(x,T_M(u_n),\nabla T_M(u_n)),\nabla T_{2k}(u_n-T_h(u_n)+T_k(u_n)-T_k(u))\big\rangle\\&\qquad\ge-\big\langle \mathcal{A}(x,T_M(u_n),\nabla T_M(u_n),\nabla T_k(u)\big\rangle\\&\qquad\ge-\big|\mathcal{A}(x,T_M(x),\nabla T_M(x))\big|\big|\nabla T_k(u)\big|.\end{aligned}$$

Therefore we obtain in (12)

$$\begin{aligned}
&\int_\Omega \omega\big\langle \mathcal{A}(x, T_M(u_n), \nabla T_M(u_n)), \nabla T_{2k}(u_n - T_h(u_n) + T_k(u_n) - T_k(u))\big\rangle \,\mathrm{d}x \\
&= \int_{\{|u_n|\le k\}} \omega\big\langle \mathcal{A}(x, T_M(u_n), \nabla T_M(u_n)), \nabla T_{2k}(u_n - T_h(u_n) + T_k(u_n) - T_k(u))\big\rangle \,\mathrm{d}x \\
&\quad + \int_{\{|u_n|> k\}} \omega\big\langle \mathcal{A}(x, T_M(u_n), \nabla T_M(u_n)), \nabla T_{2k}(u_n - T_h(u_n) + T_k(u_n) - T_k(u))\big\rangle \,\mathrm{d}x \\
&\ge \int_\Omega \omega\big\langle \mathcal{A}(x, T_k(u_n), \nabla T_k(u_n)), \nabla (T_k(u_n) - T_k(u))\big\rangle \,\mathrm{d}x \\
&\quad - \int_{\{|u_n|> k\}} \omega\big|\mathcal{A}\big(x, T_M(u_n), \nabla T_M(u_n)\big)\big|\big|\nabla T_k(u)\big| \,\mathrm{d}x\,.
\end{aligned}$$

Hence, in (11) we obtain

$$\begin{aligned}
&\int_\Omega \omega\big\langle \mathcal{A}(x, T_k(u_n), \nabla T_k(u_n)) - \mathcal{A}(x, T_k(u_n), \nabla T_k(u)), \nabla (T_k(u_n) - T_k(u))\big\rangle \,\mathrm{d}x \\
&\le \int_{\{|u_n|> k\}} \omega|\mathcal{A}(x, T_M(u_n), \nabla T_M(u_n))||\nabla T_k(u)| \,\mathrm{d}x \\
&\quad + \int_\Omega f_n T_{2k}(u_n - T_h(u_n) + T_k(u_n) - T_k(u)) \,\mathrm{d}x \\
&\quad + \int_\Omega \big\langle G_n, \nabla T_{2k}(u_n - T_h(u_n) + T_k(u_n) - T_k(u))\big\rangle \,\mathrm{d}x \\
&\quad - \int_\Omega \omega\big\langle \mathcal{A}(x, T_k(u_n), \nabla T_k(u)), \nabla (T_k(u_n) - T_k(u))\big\rangle \,\mathrm{d}x\,. \qquad (13)
\end{aligned}$$

Considering the test function $\psi_n = T_{2k}(u_n - T_h(u_n))$ in (5), we have

$$\int_\Omega \omega\,\langle \mathcal{A}(x, u_n, \nabla u_n), \nabla \psi_n\rangle \,\mathrm{d}x = \int_\Omega f_n \psi_n \,\mathrm{d}x + \int_\Omega \langle G_n, \nabla \psi_n\rangle \,\mathrm{d}x\,,$$

and by (7) we obtain

$$\int_\Omega |\nabla T_{2k}(u_n - T_h(u_n))|^p \omega \,\mathrm{d}x \le C_2(2k+1), \text{ for all } k \ge 1\,.$$

Now using that $T_{2k}(u_n - T_h(u_n)) \rightharpoonup T_{2k}(u - T_h(u))$ weakly in $W_0^{1,p}(\Omega, \omega)$ (by (9) and Remark 3 (i)), we have

$$\int_\Omega |\nabla T_{2k}(u - T_h(u))|^p \omega \,\mathrm{d}x \le C_2(2k+1)\,. \qquad (14)$$

We have (by Remark 3 (i) and (ii) and (14))

$$\begin{aligned}\int_\Omega |G|\big|\nabla T_{2k}(u-T_h(u))\big|\,\mathrm{d}x &= \int_{\{h<|u|<2k+h\}} |G||\nabla u|\,\mathrm{d}x\\ &\le \bigg(\int_{\{|u|\ge h\}} |G/\omega|^{p'}\omega\,\mathrm{d}x\bigg)^{1/p'}\bigg(\int_{\{h<|u|<2k+h\}} |\nabla u|^p\omega\,\mathrm{d}x\bigg)^{1/p}\\ &= \bigg(\int_{\{|u|\ge h\}} |G/\omega|^{p'}\omega\,\mathrm{d}x\bigg)^{1/p'}\bigg(\int_\Omega |\nabla T_{2k}(u-T_h(u))|^p\omega\,\mathrm{d}x\\ &= C_3\bigg(\int_{\{|u|\ge h\}} |G/\omega|^{p'}\omega\,\mathrm{d}x\bigg)^{1/p'},\end{aligned}$$

where C_3 depends on k but not on h. Therefore we have

$$\lim_{h\to\infty}\int_\Omega \big\langle G,\nabla T_{2k}(u-T_h(u))\big\rangle\,\mathrm{d}x = 0\,.$$

We also have (by Theorem 1)

$$\begin{aligned}\int_\Omega |T_{2k}(u-T_h(u))|^p\omega\,\mathrm{d}x &\le C_\Omega\int_\Omega |\nabla T_{2k}(u-T_h(u))|^p\omega\,\mathrm{d}x\\ &\le C_\Omega\, C_2(2k+1).\end{aligned}$$

Moreover, by Lebesgue's theorem, we obtain

$$\lim_{h\to\infty}\int_\Omega f\,T_{2k}(u-T_h(u))\,\mathrm{d}x = 0\,.$$

We can fix a positive real number h_ε sufficiently large to have

$$\int_\Omega fT_{2k}(u-T_{h_\varepsilon}(u))\,\mathrm{d}x + \int_\Omega \big\langle G,\nabla T_{2k}(u-T_{h_\varepsilon}(u))\big\rangle\,\mathrm{d}x \le \varepsilon. \tag{15}$$

Considering $h=h_\varepsilon$ in (13) (and $M=M_\varepsilon=4k+h_\varepsilon$), by (H4) and (7), we have

$$\begin{aligned}\int_\Omega &\big|\mathcal{A}(x,T_M(u_n),\nabla T_M(u_n))\big|^{p'}\omega\,\mathrm{d}x\\ &\le \int_\Omega \Big(K(x)+h_1(x)|T_M(u_n)|^{p/p'}+h_2(x)|\nabla T_M(u_n)|^{p/p'}\Big)^{p'}\omega\,\mathrm{d}x\\ &\le C\bigg[\int_\Omega K^{p'}(x)\,\omega\,\mathrm{d}x+\int_\Omega h_1^{p'}(x)|T_M(u_n)|^p\omega\,\mathrm{d}x\\ &\quad+\int_\Omega h_2^{p'}(x)|\nabla T_M(u_n)|^p\omega\,\mathrm{d}x\bigg]\\ &\le C\Big(\|K\|^{p'}_{L^{p'}(\Omega,\omega)}+\|h_1\|^{p'}_{L^\infty(\Omega)}\int_\Omega |T_M(u_n)|^p\omega\,\mathrm{d}x\\ &\quad+\|h_2\|^{p'}_{L^\infty(\Omega)}\int_\Omega |\nabla T_M(u_n)|^p\omega\,\mathrm{d}x\Big)\\ &\le C\Big(\|K\|^{p'}_{L^{p'}(\Omega,\omega)}+\|h_1\|^{p'}_{L^\infty(\Omega)}M^p\mu(\Omega)+\|h_2\|^{p'}_{L^\infty(\Omega)}M\,C_2\Big),\end{aligned}$$

that is, $|\mathcal{A}(x, T_M(u_n), \nabla T_M(u_n))|$ is bounded in $L^{p'}(\Omega, \omega)$.

Moreover, $\chi_{\{|u_n|>k\}}|\nabla T_k(u)| \to 0$ in $L^p(\Omega, \omega)$ as $n \to \infty$. Therefore,

$$\lim_{n\to\infty} \int_{\{|u_n|>k\}} \big|\mathcal{A}(x, T_M(u_n), \nabla T_M(u_n))\big| \, |\nabla T_k(u)| \, \omega \, \mathrm{d}x = 0. \tag{16}$$

Furthermore, we have that

$$T_{2k}(u_n - T_h(u_n) + T_k(u_n) - T_k(u)) \rightharpoonup T_{2k}(u - T_h(u)) \,,$$

weakly in $W_0^{1,p}(\Omega, \omega)$, as $n \to \infty$.

Hence, by (9), (15) and (16), passing to the limit in (13), we have

$$\begin{aligned}\lim_{n\to\infty} \int_\Omega \big\langle \mathcal{A}(x, T_k(u_n), \nabla T_k(u_n)) - \mathcal{A}(x, T_k(u_n), \nabla T_k(u)), \nabla(T_k(u_n) - T_k(u))\big\rangle \, \omega \, \mathrm{d}x \\ \le \int_\Omega f T_{2k}(u - T_{h_\varepsilon}(u)) \, \mathrm{d}x + \int_\Omega \langle G, \nabla T_{2k}(u - T_{h_\varepsilon}(u))\rangle \, \mathrm{d}x \le \varepsilon,\end{aligned}$$

for all $\varepsilon > 0$, that is,

$$\int_\Omega \big\langle \mathcal{A}(x, T_k(u_n), \nabla T_k(u_n)) - \mathcal{A}(x, T_k(u_n), \nabla T_k(u)), \nabla(T_k(u_n) - T_k(u))\big\rangle \, \omega \to 0 \,,$$

as $n \to \infty$. Applying Lemma 3 we get

$$T_k(u_n) \to T_k(u) \tag{17}$$

strongly in $W_0^{1,p}(\Omega, \omega)$ for every $k > 0$. This convergence implies that, for every fixed $k > 0$

$$\mathcal{A}\big(x, T_k(u_n), \nabla T_k(u_n)\big) \to \mathcal{A}\big(x, T_k(u), \nabla T_k(u)\big) \tag{18}$$

in $(L^{p'}(\Omega, \omega))^N = L^{p'}(\Omega, \omega) \times \cdots \times L^{p'}(\Omega, \omega)$.

• Finally, we need to show that u is an entropy solution to Dirichlet problem (P). Let us take $\psi_n = T_k(u_n - \varphi)$ as test function in (5), with $\varphi \in W_0^{1,p}(\Omega, \omega) \cap L^\infty(\Omega)$. We obtain,

$$\int_\Omega \omega \, \big\langle \mathcal{A}(x, u_n, \nabla u_n), \nabla \psi_n \big\rangle \, \mathrm{d}x = \int_\Omega f_n \psi_n \, \mathrm{d}x + \int_\Omega \big\langle G_n, \nabla \psi_n \big\rangle \, \mathrm{d}x \,. \tag{19}$$

If $M = k + \|\varphi\|_{L^\infty(\Omega)}$ and $n > M$, we have

$$\begin{aligned}&\int_\Omega \omega \, \big\langle \mathcal{A}(x, u_n, \nabla u_n), \nabla T_k(u_n - \varphi)\big\rangle \, \mathrm{d}x \\ &\qquad = \int_\Omega \omega \, \big\langle \mathcal{A}(x, T_M(u_n), \nabla T_M(u_n)), \nabla T_k(u_n - \varphi)\big\rangle \, \mathrm{d}x \,.\end{aligned}$$

Hence, in (19) we obtain

$$\int_\Omega \omega \,\langle \mathcal{A}(x, T_M(u_n), \nabla T_M(u_n)), \nabla T_k(u_n - \varphi)\rangle \,\mathrm{d}x \\ = \int_\Omega f_n T_k(u_n - \varphi)\,\mathrm{d}x + \int_\Omega \langle G, \nabla T_k(u_n - \varphi)\rangle \,\mathrm{d}x\,. \quad (20)$$

Therefore, by (9) and (18), passing to the limit as $n \to \infty$ in (20), we obtain

$$\int_\Omega \omega \,\big\langle \mathcal{A}(x, u, \nabla u), \nabla T_k(u - \varphi)\big\rangle \,\mathrm{d}x = \int_\Omega f\, T_k(u - \varphi)\,\mathrm{d}x + \int_\Omega \big\langle G, \nabla T_k(u - \varphi)\big\rangle \,\mathrm{d}x$$

for all $\varphi \in W_0^{1,p}(\Omega, \omega) \cap L^\infty(\Omega)$ and for each $k > 0$.

Therefore u is an entropy solution of problem (P). □

Example 1. Let $\Omega = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$, the weight function

$$\omega(x, y) = (x^2 + y^2)^{-1/2} \quad (\omega \in A_3),$$
$$f(x, y) = \frac{\cos(xy)}{(x^2 + y^2)^{1/3}},$$
$$G(x, y) = \big((x^2 + y^2)\sin(xy), (x^2 + y^2)^{-1/3}\cos(xy)\big)$$

and $\mathcal{A}\colon \Omega \times \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$, $\mathcal{A}((x, y), s, \xi) = |\xi|\,\xi$. By Theorem 2, the problem

$$(P)\begin{cases} -\operatorname{div}[(x^2 + y^2)^{-1/2}\mathcal{A}(x, u, \nabla u)] = \dfrac{\cos(xy)}{(x^2 + y^2)^{1/3}} - \operatorname{div}(G(x, y)), & \text{in } \Omega \\ u(x, y) = 0, & \text{on } \partial\Omega \end{cases}$$

has an entropy solution.

References

[1] P. Bélinan, L. Boccardo, T. Gallouët, R. Gariepy, M. Pierre, J.L. Vasquez: An L^1 theory of existence and uniqueness of solutions of nonlinear elliptic equations. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* 22 (2) (1995) 241–273.

[2] L. Boccardo, F. Murat, J.P. Puel: Existence of bounded solutions for nonlinear elliptic unilateral problems. *Ann. Mat. Pura Appl.* 152 (1988) 183–196.

[3] A.C. Cavalheiro: The solvability of Dirichlet problem for a class of degenerate elliptic equations with L^1-data. *Applicable Analysis* 85 (8) (2006) 941–961.

[4] A.C. Cavalheiro: Existence of solutions for some degenerate quasilinear elliptic equations. *Le Matematiche* LXIII (II) (2008) 101–112.

[5] V. Chiadò Piat, F. Serra Cassano: Relaxation of degenerate variational integrals. *Nonlinear Anal.* 22 (1994) 409–429.

[6] E. Fabes, C. Kenig, R. Serapioni: The local regularity of solutions of degenerate elliptic equations. *Comm. PDEs* 7 (1982) 77–116.

[7] G.B. Folland: *Real Analysis.* Wiley-Interscience, New York (1984).

[8] J. Garcia-Cuerva, J.L. Rubio de Francia: *Weighted Norm Inequalities and Related Topics*. North-Holland Mathematics Studies (1985).

[9] J. Heinonen, T. Kilpeläinen, O. Martio: *Nonlinear Potential Theory of Degenerate Elliptic Equations*. Oxford Math. Monographs, Clarendon Press (1993).

[10] A. Kufner, O. John, S. Fučík: *Function Spaces*. Noordhoof International Publishing, Leyden (1977).

[11] B. Muckenhoupt: Weighted norm inequalities for the Hardy maximal function. *Trans. Am. Math. Soc.* 165 (1972) 207–226.

[12] E. Stein: *Harmonic Analysis*. Princeton University Press (1993).

[13] A. Torchinsky: *Real-Variable Methods in Harmonic Analysis*. Academic Press, San Diego (1986).

[14] B.O. Turesson: *Nonlinear Potential Theory and Weighted Sobolev Spaces*. Lecture Notes in Mathematics. Springer-Verlag (2000).

Author's address:
State University of Londrina 86057-970, Londrina, PR Brazil
E-mail: accava@gmail.com

Received: 24 February, 2014
Accepted for publication: 4 April, 2014
Communicated by: Geoff Prince

Communications in Mathematics 22 (2014) 71–76

Lower bounds for simultaneous Diophantine approximation constants[1]

Werner Georg Nowak

Abstract. After a brief exposition of the state-of-art of research on the (Euclidean) simultaneous Diophantine approximation constants θ_s, new lower bounds are deduced for θ_6 and θ_7.

1 Introduction

For a fixed positive integer s, the (Euclidean) *simultaneous Diophantine approximation constant* θ_s is defined as the supremum of all constants c such that, for every point $\mathbf{a}$ in $\mathbb{R}^s \setminus \mathbb{Q}^s$, there exist infinitely many $(s+1)$-tuples $(\mathbf{p}, q) \in \mathbb{Z}^s \times \mathbb{N}^*$ with

$$\left|\mathbf{a} - \frac{1}{q}\mathbf{p}\right| \le \frac{1}{q\sqrt[s]{cq}}\,, \tag{1}$$

where $|\cdot|$ denotes the Euclidean norm in $\mathbb{R}^s$.

This notion generalizes a question whose answer is known as Hurwitz' classic theorem. This involves the special case $s = 1$ and tells us that $\theta_1 = \sqrt{5}$; see, e.g., Niven and Zuckerman [11, p. 189 and p. 221].

By some very deep analysis, Davenport and Mahler [6] were able to prove that $\theta_2 = \frac{1}{2}\sqrt{23}$.

For $s \ge 3$, the exact values of θ_s are unknown, and only more or less precise bounds have been established.

We remark parenthetically that the problem becomes even considerably more difficult if one replaces in (1) the Euclidean norm by the maximum norm: The constants arising, say $\theta_s^{(\infty)}$, are unknown for all $s \ge 2$, the only general successful approach being due to Spohn [16] who combined the calculus of variation with a classic method of Blichfeldt [2] to estimate $\theta_s^{(\infty)}$ from below.

2010 MSC: 11J13, 11H16

Key words: Geometry of numbers, Diophantine approximation, approximation constants, critical determinant

[1] Presented at the 21st Czech and Slovak International Conference on Number Theory (September 2–6, 2013, Ostravice, Czech Republic)

2 Survey of methods and known results

The usual approach to estimate θ_s is based on tools from the *geometry of numbers*; cf. throughout the monograph by Gruber and Lekkerkerker [7].

We briefly recall a few basic concepts of this theory: For a star-body $\mathcal{B}$ in $\mathbb{R}^s$, a lattice $\Lambda = A\mathbb{Z}^s$ (A a nonsingular real $(s \times s)$-matrix) is called *admissible* if its only point in the interior of $\mathcal{B}$ is the origin. The *critical determinant* $\Delta(\mathcal{B})$ of $\mathcal{B}$ is then defined as the infimum of the *lattice constants* $d(\Lambda) = |\det A|$, taken over all lattices Λ admissible for $\mathcal{B}$.

By a celebrated result of Davenport [5] (see also [7, p. 480, Theorem 4]), θ_s is equal to the critical determinant $\Delta(\mathcal{B}_{s+1})$ of the $(s+1)$-dimensional star body of points $(x_0, x_1, \ldots, x_s) \in \mathbb{R}^{s+1}$

$$\mathcal{B}_{s+1}: \quad |x_0| \left(\sum_{j=1}^{s} x_j^2 \right)^{s/2} \le 1\,. \tag{2}$$

For $s = 3, 4, 5$, the sharpest known lower[2] estimates for θ_s have been obtained by a method which is based on inequalities relating the critical determinants of star bodies in different dimensions. In its essence it goes back to Mordell [8], [9], [10], and Armitage [1]. Combining these tools with the known critical determinant $\Delta(\mathcal{P}) = \frac{1}{2}$ of the three-dimensional double paraboloid

$$\mathcal{P}: \quad x^2 + y^2 + |z| \le 1\,, \tag{3}$$

it has been proved [13], [14] that

$$\theta_3 \ge 1.879\ldots, \quad \theta_4 \ge 1.3225\ldots, \quad \theta_5 \ge 0.876\,. \tag{4}$$

For $s \le 5$, it appears that this is the very limit of present methods. For $s \ge 6$, the only successful approach is due to Prasad [15]. This is based on the simple idea to apply the arithmetic-geometric mean inequality to the left-hand side of (2). In terms of geometry, this amounts to inscribing an $(s+1)$-dimensional ellipsoid

$$\mathcal{E}_{s+1}: \quad \frac{1}{s+1} x_0^2 + \frac{s}{s+1} \sum_{j=1}^{s} x_j^2 \le 1 \tag{5}$$

into $\mathcal{B}_{s+1}$. It follows that

$$\theta_s = \Delta(\mathcal{B}_{s+1}) \ge \Delta(\mathcal{E}_{s+1}) = \frac{(s+1)^{(s+1)/2}}{s^{s/2}} \Delta(\mathcal{S}_{s+1}) \tag{6}$$

where $\mathcal{S}_{s+1}$ is the unit sphere in $\mathbb{R}^{s+1}$. Now the critical determinants of the unit spheres are known up to dimension 8: See [7, p. 410]; in particular, $\Delta(\mathcal{S}_7) = \frac{1}{8}$, $\Delta(\mathcal{S}_8) = \frac{1}{16}$. Hence, using (6), it readily follows that

$$\theta_6 \ge \frac{343}{1728}\sqrt{7}\,, \quad \theta_7 \ge \frac{256}{343}\frac{1}{\sqrt{7}}\,. \tag{7}$$

[2]Obviously, *lower* bounds are the more interesting ones, since they guarantee, for every c less than the bound, the existence of infinitely many solutions of the inequality (1).

We conclude this section by the remark that the question of *upper* bounds for θ_s has been dealt with in [14, section 4]. It amounts to finding certain number fields of degree $s+1$ with small absolute discriminant.

3 Improvement of the estimate (7)

The critical determinants $\Delta(\mathcal{S}_7)$, $\Delta(\mathcal{S}_8)$ once known, the deduction of the lower bounds (7) seems so natural that one might believe that this could be the end-of--the-art for this problem in the cases $s = 6, 7$. In this little note, however, we will establish a slight refinement.

Theorem 1. *The inequalities*

$$\theta_6 \geq \frac{343}{1728}\sqrt{7}\,(1+\omega_6)\,, \qquad \theta_7 \geq \frac{256}{343}\frac{1}{\sqrt{7}}\,(1+\omega_7)$$

hold true, with certain small constants $\omega_6 > 9 \times 10^{-4}$, $\omega_7 > 3 \times 10^{-4}$. *I.e., numerically,* $\theta_6 \geq 0.52564$, $\theta_7 \geq 0.28218$.

Of course, this improvement is fairly small, the main interest lying in the method applied. This in turn is inspired by classic work due to Davenport [3], [4], and Žilinskas [17], as well as by an earlier article by the author [12][3].

4 Proof of the theorem

For better readability, we give the details only for $s = 6$, the case $s = 7$ being completely analogous. In principle, the argument can be extended to $s > 7$ as well, but this is of less importance, since $\Delta(\mathcal{S}_{s+1})$ is known for $s \leq 7$ only.

The star body $\mathcal{B}_7$ defined in (2) is *automorphic*, hence there exists a *critical lattice*[4] Λ with a point on the boundary of $\mathcal{B}_7$; cf. [7, p. 305, Theorem 4]. Applying to Λ a suitable automorphism of $\mathcal{B}_7$, if necessary, we can assume this point to be $\mathbf{e} = (1,1,0,0,0,0,0)$. With $\mathbf{x} \in \mathbb{R}^7$, the function

$$G(\mathbf{x}) := \left(\frac{1}{7}x_0^2 + \frac{6}{7}\sum_{j=1}^{6} x_j^2\right)^{1/2},$$

which is the square-root of the left-hand side of (5) for $s = 6$, is called the *distance function* of the ellipsoid $\mathcal{E}_7$; it is homogeneous of order 1. Since, according to [7, p. 195, Theorem 3], any $\mathbf{o}$-symmetric ellipsoid has anomaly 1, there exist seven linearly independent lattice points $\mathbf{u}^{(k)}$ of Λ, with $\big(G(\mathbf{u}^{(k)})\big)_{k=1}^{7}$ nondecreasing, and

$$\Delta(\mathcal{E}_7)\prod_{k=1}^{7} G(\mathbf{u}^{(k)}) \leq d(\Lambda) = \Delta(\mathcal{B}_7) = \theta_6\,. \tag{8}$$

[3] Carrying out the numerical details on the basis of the argument developed in that latter paper, one would get only $\omega_6 > 6 \times 10^{-5}$, $\omega_7 > 1.5 \times 10^{-5}$.

[4] I.e., Λ is admissible for $\mathcal{B}_7$, and $d(\Lambda) = \Delta(\mathcal{B}_7)$.

We pick $\mathbf{u} = (u_0, u_1, \dots, u_6) \in \{\mathbf{u}^{(1)}, \mathbf{u}^{(2)}\}$ in such a way that $\mathbf{u} \neq \pm\mathbf{e}$. Then $\mathbf{u} \pm \mathbf{e}$ are nontrivial lattice points of Λ. Since Λ is admissible for $\mathcal{B}_7$, a look back to (2) shows that

$$|u_0| \left(\sum_{j=1}^{6} u_j^2 \right)^3 \geq 1 , \qquad |u_0 \pm 1| \left((u_1 \pm 1)^2 + \sum_{j=2}^{6} u_j^2 \right)^3 \geq 1 . \tag{9}$$

Since $\mathcal{E}_7 \subset \mathcal{B}_7$, it follows that $G(\mathbf{u}^{(1)}) \geq 1$, hence (8) implies that

$$\frac{\theta_6}{\Delta(\mathcal{E}_7)} \geq (G(\mathbf{u}))^6 . \tag{10}$$

To prove the Theorem, it remains to minimize $G(\mathbf{u})$ under the constraints (9). We put $S = \sum_{j=2}^{6} u_j^2$ for short, and may assume, w.l.o.g., that $u_0 > 0$. Hence we have to deal with a minimization problem in three variables only, namely u_0, u_1 and S. In fact,

$$M := \min_{(9)} G^2(\mathbf{u}) = \min_{(12)} \left(\frac{1}{7} u_0^2 + \frac{6}{7} \left(u_1^2 + S \right) \right) , \tag{11}$$

with

$$u_0 \left(u_1^2 + S \right)^3 \geq 1 , \qquad |u_0 \pm 1| \left((u_1 \pm 1)^2 + S \right)^3 \geq 1 . \tag{12}$$

Solving (12), we infer that

$$u_1^2 + S \geq \max\left(u_0^{-1/3}, (u_0 + 1)^{-1/3} - 1 - 2u_1, |u_0 - 1|^{-1/3} - 1 + 2u_1 \right) .$$

Hence, in view of (11),

$$7M = \min_{u_0 > 0,\, u_1} \Bigg(\max\Big(u_0^2 + 6u_0^{-1/3},\ u_0^2 + 6(u_0 + 1)^{-1/3} - 6 - 12u_1, \\ u_0^2 + 6\,|u_0 - 1|^{-1/3} - 6 + 12u_1 \Big) \Bigg) .$$

Keeping u_0 fixed for the moment and seeking the minimum with respect to u_1, we observe that the maximum of the last two expressions becomes minimal when they are equal. This obviously happens for

$$u_1 = \frac{1}{4} \left((u_0 + 1)^{-1/3} - |u_0 - 1|^{-1/3} \right) .$$

Consequently,

$$7M = \min_{u_0 > 0} \left(\max\left(u_0^2 + 6u_0^{-1/3},\ u_0^2 + 3(u_0 + 1)^{-1/3} + 3\,|u_0 - 1|^{-1/3} - 6 \right) \right) . \tag{13}$$

In order to solve this ultimate minimization problem, the figure below is very helpful.

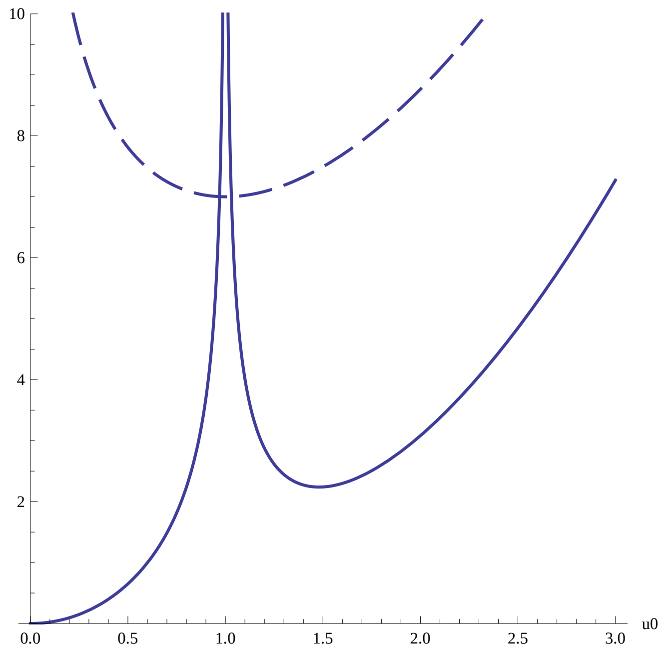

Figure 1: The graphs of $f_1\colon u_0 \mapsto u_0^2 + 6u_0^{-1/3}$ (dashed) and $f_2\colon u_0 \mapsto u_0^2 + 3(u_0+1)^{-1/3} + 3\,|u_0-1|^{-1/3} - 6$

In fact, let f_1, f_2 be defined as in the graphics, then an easy calculus exercise shows that f_1 decreases on $]0,1[$ and increases on $]1,\infty[$, and that f_2 increases on $]0,1[$. Furthermore, $f_1 - f_2$ increases on $]1,\infty[$, since there

$$\frac{\mathrm{d}}{\mathrm{d}u_0}\Big(f_1(u_0) - f_2(u_0)\Big) = \frac{1}{(u_0+1)^{4/3}} + \frac{1}{(u_0-1)^{4/3}} - \frac{2}{u_0^{4/3}} > 0\,,$$

on applying the mean inequality to the first two fractions. It readily follows that the equation $f_1(u_0) = f_2(u_0)$ has exactly two solutions, $u_0^{(1)} < 1$ and $u_0^{(2)} > 1$, say, and that, recalling (13),

$$7M = \min\Big(f_1(u_0^{(1)}), f_1(u_0^{(2)})\Big)\,.$$

Carrying out the numerics, we get $u_0^{(1)} = 0.97012\ldots$, $u_0^{(2)} = 1.030799\ldots$, hence,

$$7M = \min(7.002111\ldots, 7.00218\ldots) \geq 7.002111\,.$$

Going back to (10) and (11), we finally infer that

$$\frac{\theta_6}{\Delta(\mathcal{E}_7)} \geq \left(\frac{7.002111}{7}\right)^3 \geq 1.0009\,,$$

which completes the proof of the Theorem. $\square$

References

[1] J.V. Armitage: On a method of Mordell in the geometry of numbers. *Mathematika* 2 (1955) 132–140.

[2] H. Blichfeldt: A new principle in the geometry of numbers, with some applications. *Trans. Amer. Math. Soc.* 15 (1914) 227–235.

[3] H. Davenport: On the product of three homogeneous linear forms. *J. London Math. Soc.* 13 (1938) 139–145.

[4] H. Davenport: On the minimum of a ternary cubic form. *J. London Math. Soc.* 19 (1944) 13–18.

[5] H. Davenport: On a theorem of Furtwängler. *J. London Math. Soc.* 30 (1955) 185–195.

[6] H. Davenport, K. Mahler: Simultaneous Diophantine approximation. *Duke Math. J.* 13 (1946) 105–111.

[7] P.M. Gruber, C.G. Lekkerkerker: *Geometry of numbers.* North Holland, Amsterdam (1987).

[8] L.J. Mordell: The product of three homogeneous linear ternary forms. *J. London Math. Soc.* 17 (1942) 107–115.

[9] L.J. Mordell: Observation on the minimum of a positive quadratic form in eight variables. *J. London Math. Soc.* 19 (1944) 3–6.

[10] L.J. Mordell: On the minimum of a ternary cubic form. *J. London Math. Soc.* 19 (1944) 6–12.

[11] I. Niven, H.S. Zuckerman: *Einführung in die Zahlentheorie.* Bibliograph. Inst., Mannheim (1975).

[12] W.G. Nowak: On simultaneous Diophantine approximation. *Rend. Circ. Mat. Palermo, Ser. II* 33 (1984) 456–460.

[13] W.G. Nowak: The critical determinant of the double paraboloid and Diophantine approximation in $\mathbb{R}^3$ and $\mathbb{R}^4$. *Math. Pannonica* 10 (1999) 111–122.

[14] W.G. Nowak: Diophantine approximation in $\mathbb{R}^s$: On a method of Mordell and Armitage. In: *Algebraic number theory and Diophantine analysis. Proceedings of the conference held in Graz, Austria, August 30 to September 5, 1998.* W. de Gruyter, Berlin (2000) 339–349.

[15] A.V. Prasad: Simultaneous Diophantine approximation. *Proc. Indian Acad. Sci. A* 31 (1950) 1–15.

[16] W.G. Spohn: Blichfeldt's theorem and simultaneous Diophantine approximation. *Amer. J. Math.* 90 (1968) 885–894.

[17] G. Žilinskas: On the product of four homogeneous linear forms. *J. London Math. Soc.* 16 (1941) 27–37.

Author's address:
Institute of Mathematics, Department of Integrative Biology, BOKU Wien, 1180 Vienna, Austria

E-mail: `nowak@boku.ac.at`

Received: 9 September, 2013
Accepted for publication: 9 March, 2014
Communicated by: Olga Rossi

Communications in Mathematics 22 (2014) 77–102

Reducibility and irreducibility of Stern $(0,1)$-polynomials[1]

Karl Dilcher, Larry Ericksen

Abstract. The classical Stern sequence was extended by K. B. Stolarsky and the first author to the Stern polynomials $a(n;x)$ defined by $a(0;x)=0$, $a(1;x)=1$, $a(2n;x)=a(n;x^2)$, and $a(2n+1;x)=x\,a(n;x^2)+a(n+1;x^2)$; these polynomials are Newman polynomials, i.e., they have only 0 and 1 as coefficients. In this paper we prove numerous reducibility and irreducibility properties of these polynomials, and we show that cyclotomic polynomials play an important role as factors. We also prove several related results, such as the fact that $a(n;x)$ can only have simple zeros, and we state a few conjectures.

1 Introduction

The *Stern sequence* $\{a(n)\}_{n\ge 0}$ is defined recursively by $a(0)=0$, $a(1)=1$, and for $n\ge 1$ by

$$a(2n)=a(n),\qquad a(2n+1)=a(n)+a(n+1). \tag{1}$$

The sequence starts as $0,1,1,2,1,3,2,3,1,4,3,5,2,5,3,4,1,5,4,7,3,\ldots$ See [3] for some historical remarks and for some properties of this sequence. Perhaps the most remarkable properties are that the terms $a(n),a(n+1)$ are always relatively prime, and that each positive reduced rational number occurs once and only once in the sequence $\{a(n)/a(n+1)\}_{n\ge 1}$.

Recently the Stern sequence was extended to two different sequences of polynomials, one by the first author and K. B. Stolarsky [3], and the other independently by Klavžar, Milutinović and Petr [8]. These sequences are quite different from each other, but both have interesting and useful properties. In this paper we will only

2010 MSC: Primary: 11B83, Secondary: 11R09

Key words: Stern sequence, Stern polynomials, reducibility, irreducibility, cyclotomic polynomials, discriminants, zeros

[1]Presented at the 21st Czech and Slovak International Conference on Number Theory (September 2–6, 2013, Ostravice, Czech Republic)

consider the sequence introduced in [3]; it is defined recursively by $a(0;x) = 0$, $a(1;x) = 1$, and for $n \geq 1$ by

$$a(2n;x) = a(n;x^2)\,, \tag{2}$$

$$a(2n+1;x) = x\,a(n;x^2) + a(n+1;x^2)\,. \tag{3}$$

We call the polynomial $a(n;x)$ the n-th *Stern* $(0,1)$*-polynomial*. However, if there is no danger of confusion with the polynomials of Klavžar et al., we will simply refer to them as *Stern polynomials*, as we do in the remainder of this paper. Numerous properties of these polynomials can be found in [3] and [4]; here we only repeat the obvious properties

$$a(n;0) = 1 \quad (n \geq 1), \qquad a(n;1) = a(n) \quad (n \geq 0)\,, \tag{4}$$

where the second identity follows from comparing (2), (3) with (1). Also, for all $m \geq 0$ we have

$$a(2^m;x) = 1\,, \tag{5}$$

and the identities (2), (3) immediately give

$$a(2n+1;x) = x\,a(2n;x) + a(2n+2;x)\,. \tag{6}$$

To obtain an expression for the degree of $a(n;x)$, we let $e(n)$ denote the highest power of 2 dividing n. Then for $n \geq 1$,

$$\deg a(n;x) = \frac{n - 2^{e(n)}}{2}\,, \tag{7}$$

and in particular $\deg a(2n+1;x) = n$. Another important property of these Stern polynomials is the fact that they are $(0,1)$-polynomials, also known as *Newman polynomials*, which is not the case for the Stern polynomials of Klavžar et al. Tables of $a(n;x)$ for $1 \leq n \leq 32$ can be found in both [3] and [4], and Table 1 shows the irreducible factors (if any) of these polynomials.

While reducibility and irreducibility properties of the sequence of Klavžar et al. have been studied (see [17] and [20]), only one irreducibility result of limited scope is known for the other sequence.

Proposition 1 ([3]). *If p is a prime and 2 is a primitive root modulo p, then $a(p;x)$ is an irreducible polynomial over the rationals.*

Table 1 indicates that relatively few Stern polynomials are reducible. However, we are going to show that several infinite classes of these polynomials are in fact reducible. Other infinite classes of polynomials will be proven irreducible. Throughout this paper, reducibility and irreducibility is assumed to be over the rationals.

This paper is structured as follows. In Section 2 we are going to prove reducibility and irreducibility results for certain binomials, trinomials and quadrinomials among the Stern polynomials; part of this will be based on several known irreducibility results. In Section 3 we will prove reducibility and irreducibility for two

n	$a(n;x)$	n	$a(n;x)$
1	1	17	$(x^4+x^3+x^2+x+1)(x^4-x^3+1)$
2	1	18	$x^8+x^4+x^2+1$
3	$x+1$	19	$x^9+x^8+x^5+x^4+x^3+x+1$
4	1	20	$(x^2+x+1)(x^2-x+1)(x^4-x^2+1)$
5	x^2+x+1	21	$x^{10}+x^9+x^8+x^6+x^5+x^2+x+1$
6	x^2+1	22	$x^{10}+x^8+x^6+x^2+1$
7	x^3+x+1	23	$x^{11}+x^9+x^8+x^7+x^3+x+1$
8	1	24	x^8+1
9	x^4+x^2+x+1	25	$(x^2-x+1)(x^{10}+x^9+x^8+x^7-x^5$
10	$(x^2+x+1)(x^2-x+1)$		$-x^4+2x^2+2x+1)$
11	$x^5+x^4+x^3+x+1$	26	$x^{12}+x^{10}+x^4+x^2+1$
12	x^4+1	27	$x^{13}+x^{12}+x^{11}+x^5+x^4+x^3+x+1$
13	$x^6+x^5+x^2+x+1$	28	$x^{12}+x^4+1$
14	x^6+x^2+1	29	$x^{14}+x^{13}+x^6+x^5+x^2+x+1$
15	x^7+x^3+x+1	30	$x^{14}+x^6+x^2+1$
16	1	31	$x^{15}+x^7+x^3+x+1$

Table 1: $a(n;x)$ and their factorizations, $1 \leq n \leq 31$.

special classes of Stern polynomials with increasing numbers of terms. In Section 4 we derive several new identities for the Stern polynomials which will be used later. Section 5 is devoted to results concerning divisibility by $x^2 \pm x+1$, and in Section 6 these results are extended to more general classes of cyclotomic factors. Section 7 deals with the question of the existence of multiple factors and multiple zeros, along with some brief general remarks on the distribution of zeros of Stern polynomials. We conclude this paper with some further remarks and conjectures in Section 8.

2 Binomials, trinomials and quadrinomials

In this section we will deal with the smallest Stern polynomials, in the sense of having the least number of terms. By the second part of (4) and the fact that we are dealing with (0,1)-polynomials, the number of terms of $a(n;x)$ is just the number $a(n)$ in the Stern sequence. We can therefore use known results for this sequence. First we note that $a(n) = 1$ if and only if $n = 2^m$, $m \geq 0$. So by (5) the only monomial that can occur is the constant polynomial 1.

Next we use an observation by Lehmer [9] which essentially says that, given an integer $k \geq 2$, the number of integers n in the interval $2^{k-1} \leq n \leq 2^k$ for which $a(n) = k$ is $\varphi(k)$, where φ denotes Euler's totient function. Furthermore, it is the same number in any subsequent interval between two consecutive powers of 2. This means that there are exactly $\varphi(2)+\varphi(3)+\varphi(4) = 5$ classes of binomials, trinomials and quadrinomials. Their smallest elements (by degree) can be found in Table 1, and all others are generated from them by (2). These classes are

$$a(3;x) = x+1\,, \qquad a(3\cdot 2^k;x) = x^{2^k}+1\,; \tag{8}$$

$$a(5;x) = x^2+x+1\,, \qquad a(5\cdot 2^k;x) = x^{2^{k+1}}+x^{2^k}+1\,; \tag{9}$$

$$a(7;x) = x^3+x+1\,, \qquad a(7\cdot 2^k;x) = x^{3\cdot 2^k}+x^{2^k}+1\,; \tag{10}$$

$$a(9;x) = x^4+x^2+x+1\,, \qquad a(9\cdot 2^k;x) = x^{2^{k+2}}+x^{2^{k+1}}+x^{2^k}+1\,; \tag{11}$$

$$a(15;x) = x^7+x^3+x+1\,, \qquad a(15\cdot 2^k;x) = x^{7\cdot 2^k}+x^{3\cdot 2^k}+x^{2^k}+1\,. \tag{12}$$

We deal with these classes in sequence.

Proposition 2. *The polynomials $a(3\cdot 2^k;x)$ are irreducible for all $k\geq 0$.*

Proof. There are two ways of proving this. First, it is known that $x^{2^k}+1$ is the cyclotomic polynomial $\Phi_{2^{k+1}}(x)$, and as such it is irreducible. Second, it is easy to see that all coefficients of $a(3\cdot 2^k;x+1)$ are even, except for the leading coefficient 1, and that the constant coefficient is 2. This shifted polynomial is therefore 2-Eisenstein for any k, and is thus irreducible. □

Proposition 3. *We have $x^2+x+1 \mid a(5\cdot 2^k;x)$ for all $k\geq 0$. In other words, $a(5;x) = x^2+x+1$ is the only irreducible polynomial in this class.*

Proof. Using the factorization

$$x^4+x^2+1 = (x^2+x+1)(x^2-x+1) \tag{13}$$

with x replaced by x^{2^k}, we get

$$a(5\cdot 2^{k+1};x) = \left(x^{2^{k+1}}-x^{2^k}+1\right)a(5\cdot 2^k;x)\,.$$

The result now follows immediately by induction. □

Proposition 3 could also be obtained from a short paper by Tuckerman [19]. The situation for the trinomials in (10) is quite different, as the following result shows.

Proposition 4. *The polynomials $a(7\cdot 2^k;x)$ are irreducible for all $k\geq 0$.*

For $k=0$, the polynomial is easily verified to be irreducible; it is also a special case of Theorem 1 in [18]. For the cases $k\geq 1$ we apply Theorem 4 in [18] which we state as a lemma.

Lemma 1 (Selmer). *If $f(x) = x^n+ax^m+b$ $(m<n)$ is an irreducible trinomial satisfying one of the two sets of conditions*

$$2^3\nmid a, 2\nmid b, n\neq 2m\,, \qquad \textit{or} \qquad a\equiv 1 \textit{ or } 2 \pmod 4, 2\mid b\,, \tag{14}$$

then $f(x^2)$ is also irreducible.

We now prove the proposition by induction on k. The base case $k = 0$ has already been established; assume now that $a(7 \cdot 2^k; x)$ is irreducible for some $k \geq 0$. By (2) we have $a(7 \cdot 2^{k+1}; x) = a(7 \cdot 2^k; x^2)$, and by (10) the first condition in (14) is clearly satisfied. Hence $a(7 \cdot 2^{k+1}; x)$ is also irreducible, and we are done.

Proposition 5. *The polynomials $a(9 \cdot 2^k; x)$ and $a(15 \cdot 2^k; x)$ are irreducible for all $k \geq 0$. In other words, all Stern quadrinomials are irreducible.*

This is a fairly easy consequence of a result of Finch and Jones [7] which we quote as another lemma.

Lemma 2 (Finch and Jones). *The polynomial $x^a + x^b + x^c + 1$ is reducible if and only if exactly one of the integers $a/2^\nu, b/2^\nu, c/2^\nu$ is even, where $\gcd(a, b, c) = 2^\nu m$ with m odd.*

We apply this lemma to the polynomials in (11) and (12), noting that in both cases we have $\nu = k$. In the first case the three quotients in the lemma are 4, 2, 1, while in the second case they are 7, 3, 1. This proves irreducibility in all cases.

In closing we mention that the irreducibility of $a(15 \cdot 2^k; x)$ could also be obtained from a result of Ljunggren [10, p. 69] (corrected in [13]).

3 Stern polynomials with index $2^k \pm 1$

In this section we deal with two further classes of Stern polynomials for which we can obtain some interesting reducibility results. In contrast to the polynomials in the previous section, these have increasing numbers of terms. In [3] it was shown that

$$a(2^k - 1; x) = 1 + x + x^3 + x^7 + \cdots + x^{2^{k-1}-1}, \tag{15}$$

$$a(2^k + 1; x) = 1 + x + x^2 + x^4 + \cdots + x^{2^{k-1}}; \tag{16}$$

both follow quite easily, by induction, from (2), (3) and (5).

3.1 The polynomials $a(2^k - 1; x)$

This special case is essentially the same as a sequence of polynomials $P_n(x)$ which was studied by K. Mahler [11]:

$$x \cdot a(2^n - 1; x) = P_n(x) := \sum_{j=0}^{n-1} x^{2^j}. \tag{17}$$

In particular, Mahler studied divisibility of the polynomials $P_n(x)$ by cyclotomic polynomials $\Phi_k(x)$. He used a classical result by L. Fuchs (1863) which gives a complete characterization of the pairs (n, k) for which $\Phi_k(x) \mid P_n(x)$. Using a different approach, in [1] we gave a more explicit version of the result of Mahler and Fuchs. In what follows we give a brief summary of these results, using the notation of Stern polynomials, via (17).

The multiplicative order of 2 modulo an integer m, which we denote by $t(m)$ following Mahler and others, plays an important role in most of these results. We recall that $t(m)$ is the smallest positive integer t for which $2^t \equiv 1 \pmod{m}$. The main result in [1] can now be stated as follows.

Proposition 6. *Let $p \ge 3$ be a fixed prime. Then for all $m \ge 1$ we have*

$$\Phi_p(x^{2^{t(p)m}-1}) \bigm| a(2^{t(p)pm}-1;x)\,. \tag{18}$$

As an illustration we state the smallest case, $p=3$. Since $\Phi_3(x)=x^2+x+1$ and $t(3)=2$, we immediately get, for all $m\ge 1$,

$$x^{2^{2m+1}-2}+x^{2^{2m}-1}+1 \bigm| a(2^{6m}-1;x)\,.$$

Using standard properties of cyclotomic polynomials, we can factor the left-hand side of (18), which gives the following more explicit expression.

Corollary 1. *Let n be such that $n=t(p)pm$ for some prime $p\ge 3$ and integer $m\ge 1$. Then $a(2^n-1;x)$ is divisible by all $\Phi_d(x)$ with $d \mid p(2^{t(p)m}-1)$ and $p^{u(p)+1}\mid d$, where $u(p)$ is the highest power of p dividing $2^{t(p)m}-1$.*

This, together with the following result, gives all cyclotomic factors of the polynomials $a(2^n-1;x)$.

Proposition 7. *If $\Phi_k(x)\mid a(2^n-1;x)$ for some n, then $\Phi_k(x)\mid a(2^{nm}-1;x)$ for all integers $m\ge 1$.*

The fact that these results give all cyclotomic factors of all $a(2^n-1;x)$ was proved in [1], using Mahler's and Fuchs's results. See [1] also for a table of admissible pairs (n,k). To conclude our discussion of the polynomials $a(2^n-1;x)$, we state the following consequence of the above results; see [1], where a more general version is obtained.

Corollary 2. *For all $m\ge 1$ we have*

$$\prod_{\substack{3^k\mid 6m\\ k\ge 1}} \Phi_{3^{k+1}}(x) \bigm| a(2^{6m}-1;x)\,.$$

In particular, Stern polynomials can have arbitrarily many irreducible factors.

Finally, some further remarks on the polynomials $a(2^n-1;x)$ can be found in Subsection 8.4.

3.2 The polynomials $a(2^k+1;x)$

Computations show that these polynomials, as given in (16), are divisible by $1+x+\cdots+x^k$ when $k=1,2,4,10,12,18,28,\ldots$ At first sight there seems to be no pattern to this sequence, but we quickly notice that $k+1$ is always a prime. Furthermore, based on our knowledge of Proposition 1, we were able to identify this sequence of primes. When $k=1$ or 2, the divisibility observation holds trivially since $a(2^1+1;x)=1+x$ and $a(2^2+1;x)=1+x+x^2$. For $k\ge 2$ we can prove the following result, where we set $k+1=p$, an odd prime.

Proposition 8. *Let $p \geq 3$ be a prime which has 2 as a primitive root. Then*

$$(1 + x + x^2 + \cdots + x^{p-1}) \bigm| a(2^{p-1} + 1; x).$$

In particular, if $p \geq 5$ is such a prime, then $a(2^{p-1} + 1; x)$ is reducible.

Proof. If 2 is a primitive root of p, then $2^0, 2^1, \ldots, 2^{p-2}$ is a reordering of $1, 2, \ldots,$ $p - 1 \pmod{p}$. Therefore with (16) we have

$$a(2^{p-1} + 1; x) = 1 + \sum_{j=0}^{p-2} x^{2^j} \equiv 1 + \sum_{j=1}^{p-1} x^j \pmod{x^p - 1}.$$

Since $(x^{p-1} + \cdots + x + 1)(x - 1) = x^p - 1$, this shows that $a(2^{p-1} + 1; x)$ is divisible by $x^{p-1} + \cdots + x + 1$. □

Computations indicate that those polynomials $a(2^k \pm 1; x)$ that were not proven reducible in this section seem to be irreducible. In Section 6 we will reformulate Proposition 8 in terms of cyclotomic polynomials and extend it to larger classes of Stern polynomials. See also Subsection 8.4 for another remark on the polynomials $a(2^k + 1; x)$.

4 Identities for Stern polynomials

Both the (numerical) Stern sequence and the sequence of Stern polynomials have a great deal of internal structure which manifests itself through various identities. In addition to the elementary identities (1)–(6) which all involve just one parameter, the following identities, involving two or three parameters, were obtained in [3] and [4]: For all $k \geq 0$ and $0 \leq j \leq 2^k$ we have

$$a(2^k + j; x) - x^j a(2^k - j; x) = a(j; x) \tag{19}$$

(see [3, Lemma 2.1]), and for $0 \leq k \leq n$ and $0 \leq j \leq 2^k$ we have

$$a(2^n - j; x) - a(2^k - j; x) = x^{2^k - j} a(j; x) a(2^n - 2^k; x) \tag{20}$$

(see [4, Proposition 2.1]). It is the purpose of this section to derive extensions or generalizations of these identities which will be applicable for our purposes. We prove four identities, two of which extend (19) and (20).

Proposition 9. *For all integers $k \geq 0$, $0 \leq j \leq 2^k$, and odd $m \geq 1$ we have*

$$a(m2^k + j; x) + x^j a(m2^k - j; x) = \bigl(a(j; x) + 2x^j a(2^k - j; x)\bigr)\, a(m2^k; x). \tag{21}$$

Proof. We first treat the case $j = 0$ separately; it reduces to

$$a(m2^k; x) + a(m2^k; x) = \bigl(a(0; x) + 2a(2^k; x)\bigr)\, a(m2^k; x),$$

which is trivially true by (5) and the fact that $a(0; x) = 0$. Assuming now that $j \geq 1$, we prove (21) by induction on k. The base case $k = 0$, $j = 1$ reduces to

$$a(m + 1; x) + xa(m - 1; x) = (a(1; x) + 2xa(0; x))\, a(m; x).$$

We note that the expression in large parentheses on the right is identically 1, and since m is odd, this identity is therefore equivalent to (6).

Now assume that (21) holds for some $k-1$ and all $0 \le j \le 2^{k-1}$. For the induction step we first let j be even, say $j = 2\ell$, with $0 \le \ell \le 2^{k-1}$. Then using the reduction formula (2), the identity (21), with $j = 2\ell$, becomes

$$\begin{aligned} a(m2^{k-1}+\ell;x^2) &+ (x^2)^\ell a(m2^{k-1}-\ell;x^2) \\ &= \big(a(\ell;x^2)+2(x^2)^\ell a(2^{k-1}-\ell;x^2)\big)\, a(m2^{k-1};x^2)\,. \end{aligned} \tag{22}$$

This identity holds by the induction hypothesis, with x replaced by x^2. Second, we let j be odd, say $j = 2\ell+1$. Then the polynomials $a(m2^k+j;x)$, $a(m2^k-j;x)$, $a(j;x)$, $a(2^k-j;x)$ in (21) have odd index; hence we use (6) to rewrite (21) as follows:

$$\begin{aligned} xa(m2^k+2\ell;x) &+ a(m2^k+2\ell+2;x) \\ &\quad + x^{2\ell+1}\big[xa(m2^k-2\ell-2;x)+a(m2^k-2\ell;x)\big] \\ &= [xa(2\ell;x)+a(2\ell+2;x) \\ &\quad +2x^{2\ell+1}\big(xa(2^k-2\ell-2;x)+a(2^k-2\ell;x)\big)\big]\, a(m2^k;x). \end{aligned}$$

This holds when both the following identities hold:

$$\begin{aligned} xa(m2^k+2\ell;x) &+ x^{2\ell+1}a(m2^k-2\ell;x) \\ &= \big[xa(2\ell;x)+2x^{2\ell+1}a(2^k-2\ell;x)\big]\, a(m2^k;x)\,, \end{aligned} \tag{23}$$

$$\begin{aligned} a(m2^k+2\ell+2;x) &+ x^{2\ell+2}a(m2^k-2\ell-2;x) \\ &= \big[a(2\ell+2;x)+2x^{2\ell+2}a(2^k-2\ell-2;x)\big]\, a(m2^k;x)\,. \end{aligned} \tag{24}$$

To deal with (23), we divide both sides by x and use the reduction formula (2). This gives us (22) which holds by the induction hypothesis. Similarly, we use (2) for all terms in (24), which gives (22) with ℓ replaced by $\ell+1$. This is still valid by the induction hypothesis since $j = 2\ell+1 \le 2^k$ implies $2\ell \le 2^k-2$, and thus $\ell \le 2^{k-1}-1$ and $\ell+1 \le 2^{k-1}$, as required. Both (23) and (24) are therefore true, which completes the proof by induction. □

Proposition 10. *For all integers $m \ge 1$, $k \ge 0$, and $0 \le j \le 2^k$ we have*

$$a(m2^k-j;x) - a(2^k-j;x)a(m2^k;x) = x^{2^k-j}a(j;x)a((m-1)2^k;x)\,. \tag{25}$$

Proof. As in the previous proof we treat the case $j = 0$ separately; the identity (25) then reduces to

$$a(m2^k;x) - a(2^k;x)a(m2^k;x) = x^{2^k}a(0;x)a((m-1)2^k;x)\,,$$

and with (5) and $a(0;x) = 0$ we see that both sides are identically 0. Now assume that $j \ge 1$; we prove (25) by induction on k. The base case $k = 0$, $j = 1$ is

$$a(m-1;x) - a(0;x)a(m;x) = x^0a(1;x)a((m-1);x)\,,$$

and we see that both sides are identical. (Note that, in contrast to Proposition 9, we do not require m to be odd).

Next, assume that (25) holds for some $k-1$ and all $0 \le j \le 2^{k-1}$. For the induction step we proceed exactly as in the proof of Proposition 9, distinguishing between the cases j even and j odd. We leave the details to the reader. □

Remark. The case $m = 1$ in (25) is trivially true, as both sides are identically 0. When $m = 2^{n-k}$, then (25) implies (20) if we take (5) into account.

We obtain the following consequence from (25) if we replace j by $2^k - j$ and then m by $m+1$.

Corollary 3. *For all integers $m \ge 0$, $k \ge 0$, and $0 \le j \le 2^k$ we have*

$$a(m2^k + j; x) - x^j a(2^k - j; x) a(m2^k; x) = a(j; x) a((m+1)2^k; x)\,. \tag{26}$$

We note that (26) with $m = 1$ gives (19), once again taking (5) into account. The final result in this section is of a slightly different nature, but will also be needed later.

Proposition 11. *Let $k \ge 0$ be an integer and $m \ge 1$ an odd integer. Then*

$$2a((m+1)2^k; x) - a(m2^k; x) = a(m; -x^{2^k})\,. \tag{27}$$

Proof. By (2) we have $a(2n; -x) = a(n; (-x)^2) = a(n; x^2) = a(2n; x)$, and therefore (6) gives

$$a(2n+1; -x) = a(2n+2; x) - x\,a(2n; x) \tag{28}$$

We now iterate the reduction formula (2):

$$\begin{aligned} 2a((m+1)2^k; x) - a(m2^k; x) &= a(m+1; x^{2^k}) + \Big(a(m+1; x^{2^k}) - a(m; x^{2^k})\Big) \\ &= a(m+1; x^{2^k}) - x^{2^k} a(m-1; x^{2^k}) = a(m; -x^{2^k})\,, \end{aligned}$$

where we have first used (6) and then (28), keeping in mind that m is odd. This completes the proof of (27). □

5 Divisibility by $x^2 \pm x + 1$

In the next section we will show that each of the "allowable" polynomials $1 + x + x^2 + \cdots + x^{p-1}$ in Proposition 8, as well as their alternating analogues $1 - x + x^2 - \cdots + x^{p-1}$, divide infinite classes of the Stern polynomials $a(n; x)$. Since the largest proportion of reducible Stern polynomials are divisible by $x^2 \pm x + 1$, we will treat this case separately, and in greater detail, including relevant tables. This is also done to illustrate the methods of proof, which rely on a repeated use of the identities from the previous section.

5.1 Divisibility by $x^2 + x + 1$

We begin with a general observation. If $a(n;x)$ is reducible for some n then, by (2), $a(2n;x)$ is also reducible. But we can say more: If $x^2+x+1 \mid a(n;x)$, then by the factorization (13) we also have $x^2+x+1 \mid a(2n;x)$. We may therefore restrict our attention to *odd* indices n.

We observe by computation that x^2+x+1 divides $a(n;x)$ only when n is divisible by 5. We therefore consider $a(5\nu;x)$ with ν odd. The first few such ν for which $x^2+x+1 \mid a(5\nu;x)$ are

$$\nu = 1, 7, 9, 15, 17, 21, 31, 33, 55, 57, 63, 65, 71, 73, 107, 111, 113, \dots$$

We see that most come in pairs $(\nu, \nu+2)$, while a few are "isolated", such as 21 and 107 (leaving aside $\nu = 1$). The first 64 of the pairs are listed in Table 2, and the first 40 isolated ν are in Table 3.

The following result, along with its corollaries, will explain all entries in Table 2 and many entries in Table 3. It is an easy consequence of identities in Section 4.

Proposition 12. *Suppose that $\mu \geq 1$ and $j \geq 1$ are such that both $a(5\mu;x)$ and $a(5j;x)$ are divisible by x^2+x+1. If k is such that $j \leq \lfloor 2^k/5 \rfloor$, then $a(5(\mu\cdot 2^k \pm j);x)$ is divisible by x^2+x+1.*

Proof. We use (25) with $m = 5\mu$ and j replaced by $5j$. Then

$$a(5(\mu\cdot 2^k - j);x) - a(2^k - 5j;x)a(5\mu 2^k;x) = x^{2^k-5j}a(5j;x)a((5\mu-1)2^k;x). \quad (29)$$

The right-hand side is divisible by x^2+x+1, by hypothesis. Similarly, the second term on the left is also divisible by x^2+x+1 since $a(5\mu 2^k;x) = a(5\mu;x^{2^k})$ is, where we have used (13). This proves the "$-$" part of the statement.

For the "+" part we use (21) with m and j as above. Noting that the right-hand side is always divisible by x^2+x+1, we see that $a(5\cdot 2^k + j;x)$ is divisible by this trinomial if and only if $a(5\cdot 2^k - j;x)$ is. Finally, using the fact that (21) and (25) hold for all $j \leq 2^k$, we see that the statement of the proposition holds for $j \leq \lfloor 2^k/5 \rfloor$. □

If we set $\mu = 1$ in Proposition 12, we immediately get the following consequence.

Corollary 4. *Let $k \geq 3$ be an integer. If x^2+x+1 divides $a(5j;x)$ for some $1 \leq j \leq \lfloor 2^k/5 \rfloor$, then x^2+x+1 also divides $a(5(2^k \pm j);x)$. In particular, for all $k \geq 3$, x^2+x+1 divides $a(5(2^k \pm 1);x)$.*

Although this result holds for all j in the given range, we will be mainly interested in odd values of j. An even j will lead to even parameters $5(2^k \pm j)$, which gives us nothing new; see the remark at the beginning of this section.

Corollary 4 explains a large number of entries in Table 2. For instance, the entries 55, 57 and 71, 73 result from $k = 6$ and $j = 7, 9$. Each new entry, in turn, leads to an infinite class of further integers j for which $a(5j;x)$ is divisible by x^2+x+1.

ν	j	ν	j	ν	j	ν	j
7, 9	1	263, 265	33	583, 585	73	1039, 1041	130
15, 17	2	271, 273	34	671, 673	84	1055, 1057	132
31, 33	4	287, 289	36	855, 857	107	1079, 1081	135
55, 57	7	335, 337	42	879, 881	110	1087, 1089	136
63, 65	8	439, 441	55	887, 889	111	1095, 1097	137
71, 73	9	447, 449	56	895, 897	112	1135, 1137	142
111, 113	14	455, 457	57	903, 905	113	1143, 1145	143
119, 121	15	479, 481	60	911, 913	114	1151, 1153	144
127, 129	16	495, 497	62	951, 953	119	1159, 1161	145
135, 137	17	503, 505	63	959, 961	120	1167, 1169	146
143, 145	18	511, 513	64	967, 969	121	1191, 1193	149
167, 169	21	519, 521	65	991, 993	124	1335, 1337	167
223, 225	28	527, 529	66	1007, 1009	126	1343, 1345	168
239, 241	30	543, 545	68	1015, 1017	127	1351, 1353	169
247, 249	31	567, 569	71	1023, 1025	128	1511, 1513	189
255, 257	32	575, 577	72	1031, 1033	129	1711, 1713	214

Table 2: Odd ν for which $x^2+x+1 \mid a(5\nu;x)$, and j such that $\nu = 8j \pm 1$.

The first paired entries in Table 2 not generated in this way are 167, 169, and then 335, 337; 439, 441; 455, 457; and a total of 16 other pairs in this table. Before explaining these entries, we list in Table 3 the first 40 "isolated" integers ν for which $a(5\nu;x)$ is divisible by x^2+x+1.

A hint towards explaining the remaining pairs of entries in Table 2 is given in the "j" columns: The values of j associated with the four pairs of entries mentioned in the previous paragraph are $j = 21, 42, 55$ and 57, respectively. We note that 21 is the smallest entry in Table 3, and 42 is twice that number, while 55 and 57 are the smallest entries in Table 2 that are not of the form $2^k \pm 1$. This is easily explained by the following consequence of Proposition 12 which is obtained by taking $j = 1$ and $k = 3$.

Corollary 5. *Suppose that the integer $\mu \geq 1$ is such that $a(5\mu;x)$ is divisible by x^2+x+1. Then $a(5(8\mu \pm 1);x)$ is also divisible by x^2+x+1.*

We note in passing that the case where μ differs by 1 from a power of 2 is already covered by Corollary 4. Also, in contrast to Corollary 4, where we could restrict ourselves to odd values of j, in Corollary 5 we must consider all allowable integer parameters μ, even or odd.

For example, starting with the smallest entry in Table 3, the values $\mu = 21, 42, 84$, and 168 each lead to a pair of entries in Table 2. In general, for each $k \geq 0$, the value $\mu = 21 \cdot 2^k$ gives a new pair of Stern polynomials divisible by x^2+x+1. In fact, every entry in Tables 2 and 3 leads to an infinite class of such Stern polynomials.

We now turn to a partial explanation of the entries in Table 3. All the entries for which the second columns are not left blank are immediately explained by

ν	$\nu =$	ν	$\nu =$	ν	$\nu =$	ν	$\nu =$
21		373		1173	$2^{10}+149$	1899	$2^{11}-149$
107	2^7-21	491	2^9-21	1213	$2^{10}+189$	1941	$2^{11}-107$
149	2^7+21	533	2^9+21	1675	$2^{11}-373$	2027	$2^{11}-21$
189		693		1699	$2^{11}-349$	2069	$2^{11}+21$
235	2^8-21	835	$2^{10}-189$	1707	$2^{11}-341$	2155	$2^{11}+107$
277	2^8+21	875	$2^{10}-149$	1723	$2^{11}-325$	2197	$2^{11}+149$
315		917	$2^{10}-107$	1733	$2^{11}-315$	2237	$2^{11}+189$
325		1003	$2^{10}-21$	1771	$2^{11}-277$	2283	$2^{11}+235$
341		1045	$2^{10}+21$	1813	$2^{11}-235$	2325	$2^{11}+277$
349		1131	$2^{10}+107$	1859	$2^{11}-189$	2363	$2^{11}+315$

Table 3: Isolated odd ν for which $x^2+x+1 \mid a(5\nu;x)$.

Corollary 4. For other cases we need Proposition 12 in its greater generality. For instance, the smallest "isolated" cases not already covered by Corollary 4 occur when $\mu = j = 21$ and $k = 7$. This shows that $a(5\cdot 2667;x)$ and $a(5\cdot 2709;x)$ are both divisible by x^2+x+1.

Remark 1. We note that in contrast to the "paired" cases (Table 2), where $\nu \equiv \pm 1 \pmod 8$, all "isolated" cases (Table 3) appear to satisfy $\nu \equiv \pm 3 \pmod 8$.

Remark 2. Numerous "isolated" cases remain unexplained, beginning with $\nu = 21$ and seven more cases indicated in Table 3 with blank entries, then followed by $\nu = 2749, 2941, 3005, 3029, 3037, 3053, 3133, 3213$. The next block of unexplained cases begins with $\nu = 4947$.

5.2 Divisibility by $x^2 - x + 1$

As in the previous subsection we observe that x^2-x+1 divides $a(n;x)$ only when n is divisible by 5. We therefore consider again $a(5\nu;x)$. In spite of many similarities to divisibility properties by x^2+x+1, there are some substantial differences. In particular, while we have the factorization (13), namely $x^4+x^2+1 = (x^2+x+1)(x^2-x+1)$, the analogous polynomial x^4-x^2+1 is irreducible. This means that $x^2-x+1 \mid a(n;x)$ does not imply $x^2-x+1 \mid a(2n;x)$. Therefore, if we want to consider divisibility by x^2-x+1, as opposed to just reducibility, we have to consider $a(5\nu;x)$ also for even ν. In this connection we have the following consequence of the factorization (13) quoted above.

Corollary 6. *If x^2+x+1 divides $a(5\nu;x)$, then x^2-x+1 divides $a(5\cdot 2^k\nu;x)$ for all $k \geq 1$.*

We may now restrict our attention to odd ν, and observe that the first few such ν for which $x^2-x+1 \mid a(5\nu;x)$ are

$$\nu = 5, 27, 37, 59, 69, 79, 81, 85, 93, 123, 133, 173, 219, 229, 251, 261, 283, 293.$$

We see that most of these come in pairs $(\nu, \nu+10)$, while others are once again "isolated", such as $\nu = 5, 85, 93$, and others again come in pairs $(\nu, \nu+2)$, such as

(79, 81). Tables 4, 5 and 6 show the first instances of each of these cases. Much of this is explained by the following result which is analogous to Proposition 12 and also follows immediately from (29).

Proposition 13. *Suppose that $\mu \geq 1$ and $j \geq 1$ are such that $x^2 + x + 1 \mid a(5\mu; x)$ and $x^2 - x + 1 \mid a(5j; x)$. If k is such that $j \leq \lfloor 2^k/5 \rfloor$, then $a(5(\mu \cdot 2^k \pm j); x)$ is divisible by $x^2 - x + 1$.*

We begin with the case illustrated by the entries in Table 4. We'll show that this case is directly related to divisibility of $a(5\nu; x)$ by $x^2 + x + 1$. Indeed, by taking $j = 5$ in Proposition 13 and recalling from Table 1 that $x^2 - x + 1 \mid a(25; x)$, we get the following result.

Corollary 7. *If $\mu \geq 1$ is such that $x^2 + x + 1 \mid a(5\mu; x)$, then $x^2 - x + 1$ divides $a(5(32\mu \pm 5); x)$.*

ν	j	ν	j	ν	j	ν	j
27, 37	1	571, 581	18	1755, 1765	55	2267, 2277	71
59, 69	2	667, 677	21	1787, 1797	56	2299, 2309	72
123, 133	4	891, 901	28	1819, 1829	57	2331, 2341	73
219, 229	7	955, 965	30	1915, 1925	60	2683, 2693	84
251, 261	8	987, 997	31	1979, 1989	62	3419, 3429	107
283, 293	9	1019, 1029	32	2011, 2021	63	3515, 3525	110
443, 453	14	1051, 1061	33	2043, 2053	64	3547, 3557	111
475, 485	15	1083, 1093	34	2075, 2085	65	3579, 3589	112
507, 517	16	1147, 1157	36	2107, 2117	66	3611, 3621	113
539, 549	17	1339, 1349	42	2171, 2181	68	3643, 3653	114

Table 4: Odd ν for which $x^2 - x + 1 \mid a(5\nu; x)$, and j such that $\nu = 32j \pm 5$.

ν	$\nu =$	ν	$\nu =$	ν	$\nu =$	ν	$\nu =$
5		931	$2^{10} - 93$	1313		1875	$2^{11} - 173$
85		939	$2^{10} - 85$	1365		1955	$2^{11} - 93$
93		1109	$2^{10} + 85$	1373		1963	$2^{11} - 85$
173		1117	$2^{10} + 93$	1397		2133	$2^{11} + 85$
419	$2^9 - 93$	1197	$2^{10} + 173$	1449		2141	$2^{11} + 93$
427	$2^9 - 85$	1247		1469		2221	$2^{11} + 173$
597	$2^9 + 85$	1259		1493		2515	
605	$2^9 + 93$	1271		1501		2605	
757		1289		1517		2733	
851	$2^{10} - 173$	1301		1565		2773	

Table 5: Isolated odd ν for which $x^2 - x + 1 \mid a(5\nu; x)$.

The “isolated” cases, the first 40 of which are shown in Table 5, are partly explained by Proposition 13. For instance, the entries in the columns $(\nu =)$ correspond to $\mu = 1$. The next smallest class of examples occurs when $\mu = 7$ and $k = 9$, in which case $x^2 - x + 1 \mid a(5\nu; x)$ for $\nu = 3491$, 3499, 3669, 3677.

We now turn to the third case, which corresponds to $x^2 - x + 1 \mid a(5\nu; x)$ for pairs $(\nu, \nu + 2)$; see Table 6. This case is fully explained by the following result.

Proposition 14. *If $j \geq 1$ is an odd integer such that $x^2 - x + 1 \mid a(5j; x)$, then $x^2 - x + 1$ divides $a(5(16j \pm 1); x)$.*

Proof. We subtract (26) from (21) and set $j = 5$, $k = 4$, and replace m by $5j$, which gives

$$x^5 a(5(16j-1); x) - \big(a(5; x) + 2x^5 a(11; x)\big)\, a(80j; x) \\ = -x^5 a(11; x) a(80j; x) - a(5; x) a(80j + 16; x)\,. \qquad (30)$$

We are going to use the fact that

$$a(5; x) + 2x^5 a(11; x) \\ = (x^2 - x + 1)(2x^8 + 4x^7 + 4x^6 - 2x^4 + 2x^2 + 2x + 1)\,, \qquad (31)$$

which is easy to verify by direct computation. Multiplying the left-hand side of (31) by $a(80j + 16; x)$ and adding the product to both sides of (30), we get

$$x^5 a(5(16j-1); x) + \big(a(5; x) + 2x^5 a(11; x)\big)\, \big(a(80j + 16; x) - a(80j; x)\big) \\ = x^5 a(11; x)\, \big(2a(80j + 16; x) - a(80j; x)\big)\,. \qquad (32)$$

Using (27) with $k = 4$ and $m = 5j$, we get

$$2a(80j + 16; x) - a(80j; x) = a(5j; -x^{16})\,. \qquad (33)$$

Now, if $x^2 - x + 1 \mid a(5j; x)$, then $x^{32} + x^{16} + 1 \mid a(5j; -x^{16})$, and by iterating the factorization (13) we see that $x^2 - x + 1 \mid x^{32} + x^{16} + 1$. Hence, with (33) we see that $x^2 - x + 1$ divides the right-hand side of (32), and (31) shows that it also divides the second term on the left-hand side of (32). This proves the “$-$” part of the proposition.

To prove the “+” part, we use (21) with j, k and m as above, obtaining

$$a(5(16j+1); x) + x^5 a(5(16j-1); x) = \big(a(5; x) + 2x^5 a(11; x)\big)\, a(80j; x)\,.$$

By (31) the right-hand side of this last identity is divisible by $x^2 - x + 1$, and since the second term on the left is also divisible by $x^2 - x + 1$, then so is the first term. This completes the proof. □

ν	j	ν	j	ν	j	ν	j
79, 81	5	1103, 1105	69	1487, 1489	93	3503, 3505	219
431, 433	27	1263, 1265	79	1967, 1969	123	3663, 3665	229
591, 593	37	1295, 1297	81	2127, 2129	133	4015, 4017	251
943, 945	59	1359, 1361	85	2767, 2769	173	4175, 4177	261

Table 6: Odd ν for which $x^2 - x + 1 \mid a(5\nu; x)$, and j such that $\nu = 16j \pm 1$.

6 Divisibility by certain cyclotomic polynomials

Computations show that numerous Stern polynomials $a(n; x)$ are divisible by $x^4 + x^3 + x^2 + x + 1$ or by $x^4 - x^3 + x^2 - x + 1$, and it appears that this occurs only when $n = 17\nu$ for a positive integer ν. We could prove results very similar to those in Section 5; however, both these cases and divisibility properties by $x^2 \pm x + 1$ are in fact special cases of a much more general result, as we shall see in this section.

Returning to Proposition 8, we note that for an odd prime p we have

$$\Phi_p(x) = 1 + x + x^2 + \cdots + x^{p-1},$$

the p-th cyclotomic polynomial. Furthermore, for ease of notation we set

$$m_p := 2^{p-1} + 1.$$

We can then reformulate Proposition 8 as follows: Let p be an odd prime; then

$$\Phi_p(x) \mid a(m_p; x) \quad \text{if 2 is a primitive root of } p. \tag{34}$$

The first twenty such primes p are 3, 5, 11, 13, 19, 29, 37, 53, 59, 61, 67, 83, 101, 107, 131, 139, 149, 163, 173, 179; see also [15, A001122]. By a conjecture of Artin we would expect the existence of infinitely many such primes. We will also need the identity

$$\Phi_p(x^2) = \Phi_p(x)\Phi_p(-x), \tag{35}$$

valid for all odd primes; note that $\Phi_p(-x) = \Phi_{2p}(x)$ is also a cyclotomic polynomial. The identity (35) is a generalization of (13), which is the case $p = 3$.

6.1 Divisibility by $\Phi_p(x)$

We are now ready to state and prove the first main result of this section; it is a direct generalization of Proposition 12.

Proposition 15. *Let p be an odd prime which has 2 as a primitive root, and suppose that $\mu \geq 1$ and $j \geq 1$ are such that both $a(m_p\mu; x)$ and $a(m_pj; x)$ are divisible by $\Phi_p(x)$. If k is such that $j \leq \lfloor 2^k/m_p \rfloor$, then $a(m_p(\mu \cdot 2^k \pm j); x)$ is divisible by $\Phi_p(x)$.*

Proof. Following the proof of Proposition 12, we use (25) with $m = m_p\mu$ and replace j with m_pj. Then

$$\begin{aligned} a\big(m_p(\mu \cdot 2^k - j); x\big) &- a(2^k - m_pj; x)a(m_p\mu 2^k; x) \\ &= x^{2^k - m_pj} a(m_pj; x)a\big((m_p\mu - 1)2^k; x\big)\,. \end{aligned} \tag{36}$$

The right-hand side is divisible by $\Phi_p(x)$, by hypothesis. Also by hypothesis we have

$$\Phi_p(x^{2^k}) \mid a(m_p\mu; x^{2^k}) = a(m_p\mu 2^k; x)\,, \tag{37}$$

where we have once again used (2). By iterating (35) we then see that the terms in (37) are divisible by $\Phi_p(x)$. Altogether, this shows that the first term in (36) is divisible by $\Phi_p(x)$, which proves the "−" part of the statement of the proposition.

For the "+" part we use the identity (21) with m and j as above. It suffices to show that the right-hand side of that identity is divisible by $\Phi_p(x)$. But since the final term in (21) is the same as the right-hand side in (37), this is the case, as we have seen above. The proof is now complete. □

As in the special case $p = 3$ in Section 5, we get the following immediate consequences.

Corollary 8. *Let p be an odd prime which has 2 as a primitive root. If $\Phi_p(x)$ divides $a(m_pj; x)$ for some $1 \le j \le \lfloor 2^k/m_p \rfloor$, then also $\Phi_p(x) \mid a(m_p(2^k \pm j); x)$. In particular, if $k \ge p$, then $\Phi_p(x) \mid a(m_p(2^k \pm 1); x)$.*

The first statement of this result follows from Proposition 15 with $\mu = 1$, where we have used (34). The second statement follows from the first by setting $j = 1$ and using (34) again. Also note that the condition $2^k \ge m_p$ holds whenever $k \ge p$.

The next consequence is an extension of Corollary 5 and is simply the case $j = 1$ and $k = p$ in Proposition 15.

Corollary 9. *Let p be an odd prime which has 2 as a primitive root. Suppose that $\mu \ge 1$ is such that $a(m_p\mu; x)$ is divisible by $\Phi_p(x)$. Then $a(m_p(2^p\mu \pm 1); x)$ is also divisible by $\Phi_p(x)$.*

6.2 Divisibility by $\Phi_p(-x)$

We continue with extending the results in Section 5; we first state an easy consequence of (35) and (2).

Corollary 10. *Let p be an odd prime which has 2 as a primitive root. If $\Phi_p(x)$ divides $a(m_p\nu; x)$, then $\Phi_p(-x)$ divides $a(m_p 2^k \nu; x)$ for all $k \ge 1$.*

Next we state the relevant generalization of Proposition 13.

Proposition 16. *Let p be an odd prime which has 2 as a primitive root and suppose that $\mu \ge 1$ and $j \ge 1$ are such that $\Phi_p(x) \mid a(m_p\mu; x)$ and $\Phi_p(-x) \mid a(m_pj; x)$. If k is such that $j \le \lfloor 2^k/m_p \rfloor$, then $a(m_p(\mu 2^k \pm j); x)$ is divisible by $\Phi_p(-x)$.*

The "−" part of this statement follows immediately from (36), using Corollary 10. The "+" part follows once again from (21) and the "−" part, with m and j as before.

Recall that Corollary 7 follows directly from Proposition 13 due to the fact that $x^2 - x + 1$ divides $a(25; x)$, that is, $\Phi_p(-x) \mid a(m_p^2; x)$ for $p = 3$. In order to generalize Corollary 7, we need this fact to be true more generally, which is indeed the case.

Proposition 17. *If p is an odd prime which has 2 as a primitive root, then $\Phi_p(-x)$ divides $a(m_p^2;x)$.*

Proof. Using the definition of m_p, we note that

$$(2^{p-2}+1)2^p = 2^{2(p-1)} + 2\cdot 2^{p-1} = m_p^2 - 1\,.$$

We now apply the identity (26) with $m = 2^{p-2}+1$, $k=p$ and $j=1$, to obtain

$$a(m_p^2;x) = a((2^{p-3}+1)2^{p+1};x) + x\cdot a(2^p-1;x)a((2^{p-2}+1)2^p;x)\,. \tag{38}$$

We claim that we have the following congruences modulo the polynomial $\Phi_p(-x)$, valid for all primes $p\ge 3$ which have 2 as a primitive root:

$$a((2^{p-3}+1)2^{p+1};x) \equiv -x^2+x \pmod{\Phi_p(-x)}\,, \tag{39}$$

$$a(2^p-1;x) \equiv x^{p-2}-x^{p-3}+\cdots-x^2+x \pmod{\Phi_p(-x)}\,, \tag{40}$$

$$a((2^{p-2}+1)2^p;x) \equiv x \pmod{\Phi_p(-x)}\,. \tag{41}$$

Substituting these congruences into the right-hand side of (38), we immediately get

$$a(m_p^2;x) \equiv x^p - x^{p-1} + \cdots + x^3 - x^2 + x \pmod{\Phi_p(-x)}\,,$$

and the right-hand side is obviously divisible by $\Phi_p(-x)$. This proves the result, provided we can prove the congruences (39)–(41).

The proofs of these congruences are based on the following two fundamental facts. First, if p has 2 as a primitive root, then $2^0, 2^1, 2^2, \dots, 2^{p-2}$ is a reordering, modulo p, of $1,2,3,\dots,p-1$. Second, we have

$$x^p \equiv -1 \pmod{\Phi_p(-x)} \tag{42}$$

since $\Phi_p(-x) = (x^p+1)/(x+1)$.

To prove (39), we first note that by Fermat's little theorem we have $2^{p+1} = 4\cdot 2^{p-1} \equiv 4 \pmod p$. This, together with $2^{p+1}\equiv 0 \pmod 2$ gives $2^{p+1}\equiv 4 \pmod{2p}$ by the Chinese Remainder Theorem, and thus

$$x^{2^{p+1}} \equiv x^4 \pmod{\Phi_p(-x)}\,. \tag{43}$$

Now, iterating the identity (2) and using (16), we get

$$a((2^{p-3}+1)2^{p+1};x) = a(2^{p-3}+1;x^{2^{p+1}}) = 1 + \sum_{j=0}^{p-4}\left(x^{2^{p+1}}\right)^{2^j} \tag{44}$$

$$\equiv 1 + \sum_{j=0}^{p-4}\left(x^4\right)^{2^j} = 1 + \sum_{j=0}^{p-4}\left(x^2\right)^{2^{j+1}}$$

$$\equiv 1 + (x^2)^2 + (x^2)^3 + \cdots + (x^2)^{p-1} - (x^2)^{2^{p-2}} \pmod{\Phi_p(-x)}\,,$$

where we have taken into account the fact that $2^{j+1}, j = 0,1,\dots,p-3$, is a reordering of $2,3,\dots,p-1 \pmod p$, and that the upper limit of summation in the second line of (44) is only $p-4$. Next we note that by Fermat's little theorem we have $2^{p-2} \equiv \frac{p+1}{2} \pmod{p}$, so that

$$(x^2)^{2^{p-2}} \equiv x^{p+1} \pmod{\Phi_p(-x)}. \tag{45}$$

Using this congruence and (42), we get with (44) that

$$a((2^{p-3}+1)2^{p+1};x) \equiv 1+x^4+x^6+\cdots+x^{p-1}-x^3-x^5-\cdots-x^{p-2} \pmod{\Phi_p(-x)}.$$

Finally, subtracting $\Phi_p(-x) = 1-x+x^2-\cdots+x^{p-1}$ from the right-hand side of this last congruence, we get (39).

To prove (40), we use (15) and (42) to obtain

$$\begin{aligned} a(2^p-1;x) &= 1+\frac{1}{x}\sum_{j=0}^{p-2}\left(x^2\right)^{2^j} \\ &\equiv 1+\frac{1}{x}\left(x^2+(x^2)^2+(x^2)^3+\cdots+(x^2)^{p-1}\right) \\ &= 1+x+x^3+\cdots+x^{p-2}+x^p+x^{p+2}+\cdots+x^{p+p-3} \\ &\equiv 1+x+x^3+\cdots+x^{p-2}-1-x^2-\cdots-x^{p-3} \pmod{\Phi_p(-x)}, \end{aligned}$$

but this is just the right-hand side of (40)

Finally, to prove (41), we proceed exactly as in the proof of (39). Again (42) and a version of (43) are used; we leave the details to the reader. This completes the proof of Proposition 17. □

Using this result, we immediately obtain the following consequence of Proposition 16.

Corollary 11. *Let p be an odd prime which has 2 as a primitive root. If $\mu \geq 1$ is such that $\Phi_p(x) \mid a(m_p\mu;x)$ then $\Phi_p(-x)$ divides $a(m_p(\mu 2^k \pm m_p);x)$ whenever $k \geq 2p-1$.*

This corollary follows from Proposition 16 by setting $j = m_p$; it is then easy to see with the definition of m_p that the condition $j \leq \lfloor 2^k/m_p \rfloor$ holds whenever $k \geq 2p-1$.

Finally, as far as a generalization of Proposition 14 is concerned, we note that the proof of that result depended in an essential way on the fact that the left-hand side of (31) is divisible by x^2-x+1. For this proof to work in general, we require the following extension of (31).

Lemma 3. *If p is an odd prime which has 2 as a primitive root, then*

$$\Phi_p(-x) \mid a(m_p;x) + 2x^{m_p} a(2^{2p-2}-m_p;x). \tag{46}$$

Proof. We begin by showing that

$$a(m_p;x) = a(2^{p-1}+1;x) \equiv 2x \pmod{\Phi_p(-x)}. \tag{47}$$

Indeed, using (16) and the remarks in the proof of Proposition 17, including (45), we find that

$$\begin{aligned} a(m_p;x) &= 1 + x + \sum_{j=0}^{p-3} \left(x^2\right)^{2^j} \\ &\equiv 1 + x + x^2 + (x^2)^2 + (x^2)^3 + \cdots + (x^2)^{p-1} - (x^2)^{2^{p-2}} \\ &\equiv 1 + x + x^2 + x^4 + x^6 + \cdots + x^{p-1} + x^{p+1} + x^{p+3} + \cdots + x^{p+p-2} - x^{p+1} \\ &\equiv 1 + x + x^2 + x^4 + x^6 + \cdots + x^{p-1} - x^3 - x^5 - \cdots - x^{p-2} \\ &\equiv 2x \pmod{\Phi_p(-x)}. \end{aligned}$$

This proves (47). Next we use (25) with $k = p$, $j = m_p$, $m = 2^{p-2}$, and note that $2^p - m_p = 2^{p-1} - 1$. This gives

$$a(2^{2p-2} - m_p;x) = a(2^{p-1}-1;x) + x^{2^{p-1}-1} a(m_p;x) a(2^{2p-2} - 2^p;x). \tag{48}$$

We now evaluate, modulo $\Phi_p(-x)$, the various terms in (46) and (48). First, by Fermat's little theorem we have $2^{p-1} + 1 \equiv 2 \pmod p$ and also $2^{p-1} + 1 \equiv 1 \pmod 2$, which combines to give $m_p = 2^{p-1} + 1 \equiv p + 2 \pmod{2p}$ by the Chinese Remainder Theorem. Consequently, we get $2^{p-1} - 1 \equiv p \pmod{2p}$. Thus, by (42) we have

$$x^{m_p} \equiv -x^2 \pmod{\Phi_p(-x)}, \qquad x^{2^{p-1}-1} \equiv -1 \pmod{\Phi_p(-x)}. \tag{49}$$

Next, using (19) with $k = p - 1$ and $j = 1$, we get

$$x \cdot a(2^{p-1} - 1;x) = a(2^{p-1} + 1;x) - 1 \equiv 2x - 1 \pmod{\Phi_p(-x)}, \tag{50}$$

where we have used (47). Similarly, using (19) with $k = 2p - 2$ and $j = 2^p$, we get with (5) that

$$x^{2^p} a(2^{2p-2} - 2^p;x) = a(2^{2p-2} + 2^p;x) - 1. \tag{51}$$

Since $2^p \equiv 2 \pmod p$ and $2^p \equiv 2 \pmod 2$, the Chinese Remainder Theorem gives $2^p \equiv 2 \pmod{2p}$, and by (42) we have

$$x^{2^p} \equiv x^2 \pmod{\Phi_p(-x)}.$$

This congruence and (41), together with (51), shows that

$$x^2 a(2^{2p-2} - 2^p;x) \equiv x - 1 \pmod{\Phi_p(-x)}. \tag{52}$$

Finally, combining (47) and (48) with (49), (50) and (52), the right-hand side of (46), taken modulo $\Phi_p(-x)$, becomes

$$2x - 2[x(2x-1) - 2x(x-1)] = 0,$$

which completes the proof of (46). □

We are now ready to state the desired generalization of Proposition 14.

Proposition 18. *Let p be an odd prime which has 2 as a primitive root. If $j \geq 1$ is an odd integer such that $\Phi_p(-x)$ divides $a(m_p j; x)$, then $\Phi_p(-x)$ also divides $a(m_p(2^{2p-2}j \pm 1); x)$.*

The proof of this result is completely analogous to that of Proposition 14, with $m_3 = 5$ replaced by m_p and $k = 4$ replaced by $k = 2p - 2$. The divisibility relation (46) plays the role of the identity (31). We leave all further details to the reader. As an illustration we explicitly state the case $p = 5$.

Corollary 12. *If $j \geq 1$ is an odd integer such that $x^4 - x^3 + x^2 - x + 1$ divides $a(17j; x)$, then $x^4 - x^3 + x^2 - x + 1$ divides $a(17(256j \pm 1); x)$.*

7 Discriminants and zeros

7.1 The discriminant of $a(n; x)$

The various results on cyclotomic factors in this paper give rise to the natural question as to whether a square or a higher power of a cyclotomic polynomial can divide a Stern polynomial. In this subsection we will show that this cannot happen.

Proposition 19. *A Stern polynomial $a(n; x)$ cannot be divisible by the square of a nonconstant polynomial.*

By (2) and (4) it suffices to consider odd indices n since a square factor of $a(2n; x)$ would also be one of $a(n; x^2)$, which in turn means that $a(n; x)$ would have a square factor. Proposition 19 is now an immediate consequence of the following result since the discriminant of a polynomial vanishes if and only if the polynomial has a multiple zero. This is clear from the definition of the discriminant: Suppose we are given a polynomial

$$f(x) = a_n x^n + \cdots + a_1 x + a_0 = a_n(x - r_1) \ldots (x - r_n) \tag{53}$$

with nonzero leading coefficient a_n and not necessarily distinct zeros $r_1, \ldots, r_n$. Then the *discriminant* of $f(x)$ can be defined as

$$D_x(f(x)) := a_n^{2n-2} \prod_{i<j} (r_i - r_j)^2\,; \tag{54}$$

see, e.g., [16, p. 217]. We are now ready to state and prove a result which immediately implies Proposition 19.

Proposition 20. *The discriminant $D_x(a(2n + 1; x))$ is always an odd integer, and is therefore nonzero.*

For the proof of this we require a general fact about the discriminant of a polynomial, which we were unable to find in the literature.

Lemma 4. *Let $f(x)$ be a polynomial of degree $n \geq 1$ with $f(0) \neq 0$. Then*

$$D_x(f(x)) = D_x\left(x^n f(\tfrac{1}{x})\right). \tag{55}$$

In other words, the discriminants of a polynomial and of its reciprocal are identical.

Proof. The polynomial $x^n f(1/x)$ can be obtained from $f(x)$ by reversing the order of the coefficients. Also, if $f(x)$ is given as in (53), then the zeros of $x^n f(1/x)$ are obviously $1/r_1, \dots, 1/r_n$. Hence, with (54) we have

$$D_x\left(x^n f(\tfrac{1}{x})\right) = a_0^{2n-2}\prod_{i<j}\left(\frac{1}{r_i}-\frac{1}{r_j}\right)^2 = a_0^{2n-2}\prod_{i<j}\left(\frac{r_j-r_i}{r_ir_j}\right)^2. \tag{56}$$

By counting the number of times each r_j occurs in the first product below, we get

$$\prod_{i<j}\frac{1}{r_ir_j} = \left(\prod_{j=1}^{n}\frac{1}{r_j}\right)^{n-1} = \left(\frac{(-1)^n a_n}{a_0}\right)^{n-1},$$

where the second equality follows from (53). Rewriting this, and squaring, we get

$$a_0^{2n-2}\prod_{i<j}\frac{1}{(r_ir_j)^2} = a_n^{2n-2}.$$

Finally, this last identity, together with (56) and (54), gives the desired identity (55). □

Proof of Proposition 20. The discriminant of a polynomial f of degree n and leading coefficient a_n satisfies $D_x(f) = (-1)^{n(n-1)/2}a_n^{-1}R(f,f')$, where $R(f_1,f_2)$ is the resultant of the polynomials f_1 and f_2, which can be written as a determinant (the determinant of the Sylvester matrix) that involves sums of products of the coefficients of f_1 and f_2. Therefore, if f is a monic polynomial with integer coefficients, we have the relation

$$D_x(f(x)) \equiv D_x(g(x)) \pmod{n} \quad \text{if} \quad f(x) \equiv g(x) \pmod{n}$$

for any integer $n > 1$. Now, in [3, (6.2)] it was shown that

$$a(2n+1;x) \equiv x^n U_{2n}\left(\tfrac{1}{2\sqrt{x}}\right) \pmod{2},$$

where $U_n(x)$ is the nth Chebyshev polynomial of the second kind. This congruence, combined with identity (55), means that we are done if we can show that $D_x(U_{2n}(\sqrt{x}/2))$ is an odd integer. We are going to show more, namely

$$D_x\left(U_{2n}\left(\tfrac{\sqrt{x}}{2}\right)\right) = (2n+1)^{n-1}. \tag{57}$$

To do so, we first note that $U_{2n}(x) = 2^{2n}x^{2n} - \cdots + (-1)^n$ has only even powers of x and coefficients with alternating signs (see, e.g., [16]), and as a consequence we have $U_{2n}(\sqrt{x}/2) = x^n - \cdots + (-1)^n$. This is a monic polynomial, again with alternating coefficients which are integers, a fact that follows easily from standard properties of the $U_n(x)$. We also know that $U_{2n}(x)$ has $2n$ distinct real zeros, say $\pm\rho_j$, $j = 1,\dots,n$; the polynomial $U_{2n}(\sqrt{x}/2)$ then has the n positive real zeros $4\rho_j^2$. Now by (54) we have

$$D_x\left(U_{2n}(x)\right) = \left(2^{2n}\right)^{4n-2}\prod_{i<j}(\rho_i-\rho_j)^2(\rho_i+\rho_j)^2(-\rho_i-\rho_j)^2(-\rho_i+\rho_j)^2\prod_{j=1}^{n}(2\rho_j)^4,$$

which can be seen by ordering the zeros of $U_{2n}(x)$ as $r_1 = -\rho_1, r_2 = \rho_1, \dots,$ $r_{2n-1} = -\rho_n$, $r_{2n} = \rho_n$. Rearranging the factors, we get

$$D_x\left(U_{2n}(x)\right) = 2^{2n(4n-2)}\prod_{i<j}(\rho_i^2-\rho_j^2)^4\prod_{j=1}^{n}(2\rho_j)^4. \tag{58}$$

Now the product $\prod\rho_j^2$ is the product of all zeros of $U_{2n}(x)$, times $(-1)^n$; but the product of all the zeros is $a_0/a_{2n} = (-1)^n 2^{-2n}$, by (53). Hence, the powers of 2 will cancel in the second product on the right of (58), and this product will simply be 1. On the other hand, it is known that

$$D_x\left(U_{2n}(x)\right) = 2^{4n^2}(2n+1)^{2n-2};$$

see [16, p. 219]. By combining this with (58) and taking square roots, we obtain

$$\prod_{i<j}(\rho_i^2-\rho_j^2)^2 = 2^{2n-2n^2}(2n+1)^{n-1}. \tag{59}$$

From the definition (54) and our above observation concerning $U_{2n}(\sqrt{x}/2)$, we get

$$D_x\left(U_{2n}(\tfrac{\sqrt{x}}{2})\right) = \prod_{i<j}(4\rho_i^2-4\rho_j^2)^2 = 16^{n(n-1)/2}\prod_{i<j}(\rho_i^2-\rho_j^2)^2.$$

This identity, combined with (59), finally gives (57), and the proof is complete. □

In the special case of the polynomials $a(2^n-1;x)$, the result of Proposition 19 was obtained in [1] in two different ways, distinct from the approach given above.

7.2 Zeros of the Stern polynomials

Proposition 19 can also be seen as a result on the zeros of Stern polynomials in that it shows that there can only be simple zeros. Also, since the zeros of cyclotomic polynomials all lie on the unit circle and (at least in the case of $\Phi_p(\pm x)$) have an almost uniform angular distribution, it may be of interest to consider the distribution of all zeros of the Stern polynomials $a(n;x)$. This was in fact done in a recent paper of A. R. Vargas [21]. Among other results, Vargas showed that, given a real number ρ satisfying $0<\rho<1$, the proportion of zeros of $a(n;z)$ that lie on the annulus $1-\rho\le|z|\le 1/(1-\rho)$ approaches 1 as $n\to\infty$, and that the zeros are uniformly distributed in a certain sense. For details, see [21, Prop. 2.1]. On the other hand, it was shown in [11] and [1] that we must expect infinite subsequences of Stern polynomials which have zeros bounded away from the unit circle. In particular, this is the case for the sequence of polynomials in (17).

8 Conjectures and further remarks

We conclude this paper with a few open problems and conjectures, as well as some related results and remarks.

8.1 Stern polynomials with prime index

We begin by considering the question of irreducibility of Stern polynomials with prime index. On the one hand we have Proposition 1 which shows that we do have irreducibility for a certain important class of prime indices, but on the other hand there is the obvious counterexample $a(17;x)$; see Table 1. We have not found any other reducible Stern polynomial with prime index; therefore we propose the following conjecture.

Conjecture 1. *Let $q \geq 3$ be a prime, $q \neq 17$. Then $a(q;x)$ is irreducible.*

We verified Conjecture 1 by computation for all primes up to 16 000. This conjecture is also supported by the fact that all Stern polynomials which we found and proved to be reducible have indices that are of the form $2^{t(p)pm} - 1$ (and thus are composite) as in Proposition 6, or have indices which are multiples of $m_p = 2^{p-1} + 1$, as in Section 6. This last case includes the possibility of m_p itself being the index, as in Proposition 8. For this reason the following easy result is relevant.

Lemma 5. *The integer $m_p = 2^{p-1} + 1$ is prime for $p = 3$ and $p = 5$, and is composite for all other primes p which have 2 as a primitive root.*

This means that $m_3 = 5$ and $m_5 = 17$ are the only prime indices which occur in Proposition 8, or in any other divisibility result. While we have $x^2+x+1 = a(5;x)$, the next case, namely $x^4 + x^3 + x^2 + x + 1$, is in fact a proper divisor of $a(17;x)$. This explains the exceptionality of $q = m_5 = 17$, and the likelihood of this being the only exception. Note that $p = 3, 5$ and $q = 5, 17$ are among the first three Fermat primes, a fact that plays a role in the following proof.

Proof of Lemma 5. A necessary condition for $2^{p-1} + 1$ being prime (and thus a Fermat prime) is that $p - 1 = 2^k$ for some $k \geq 1$; i.e., p itself has to be a Fermat prime. However, the multiplicative order of 2 modulo a Fermat prime $F_k = 2^{2^k} + 1$ is 2^{k+1} since $2^{2^k} \equiv -1 \pmod{F_k}$. But $F_k - 1 = 2^{2^k} > 2k + 1$ for $k \geq 2$, so no Fermat prime F_k, $k \geq 2$, can have 2 as a primitive root. This leaves $p = F_0 = 3$ and $p = F_1 = 5$ as the only possibilities, and the proof is complete. □

8.2 Cyclotomic factors

Our second conjecture is related to the remarks following Conjecture 1.

Conjecture 2. *Let $p \geq 3$ be a prime which has 2 as a primitive root. If $\Phi_p(x)$ or $\Phi_p(-x)$ divides $a(n;x)$, then m_p divides n.*

Cyclotomic polynomials seem to be even more prevalent than this conjecture may indicate. Indeed, we propose the following.

Conjecture 3. *If a Stern polynomial is reducible, then it is divisible by a cyclotomic polynomial $\Phi_k(x)$ for some $k \geq 3$.*

8.3 Non-reciprocal parts

Much of the work in this paper has been devoted to exposing small factors of Stern polynomials. These factors all turned out to be cyclotomic (see also Conjecture 3), while the non-cyclotomic cofactors seem to be irreducible.

This observation is related to the following concepts and results. The *non-reciprocal part* of a polynomial $f(x)$ with integer coefficients is essentially $f(x)$ with its irreducible reciprocal factors removed, where a *reciprocal polynomial* $g(x)$ satisfies $g(x) = \pm x^{\deg g} g(1/x)$. In Table 1, for instance, the first factors of $a(17;x)$ and $a(25;x)$ are reciprocal polynomials, while the second factors are the non-reciprocal parts. Also, the polynomials $a(n;x)$ for $n = 3, 5, 6, 10, 12, 20$ and 24 are themselves reciprocal, and their non-reciprocal parts are therefore identically 1. We recall that cyclotomic polynomials are reciprocal as well. For further remarks on these concepts see, e.g., the introduction of [6].

Based on earlier work of Schinzel, a criterion for a polynomial of the form $f(x)x^n + g(x)$, with $f(x), g(x) \in \mathbb{Z}[x]$, to have an irreducible non-reciprocal part was established by Filaseta, Ford and Konyagin [5]. This result was considerably strengthened by Filaseta and Matthews [6] in the special case of $(0,1)$-polynomials. Although this last result fails to be applicable to the Stern polynomials, our observations suggest that the conclusion still holds:

Conjecture 4. *All Stern polynomials have a non-reciprocal part that is either irreducible or is identically* 1.

As pointed out in [6], if it is known that a polynomial $f(x)$ has an irreducible non-reciprocal part, then $f(x)$ is itself irreducible if it has no factor in common with its reciprocal $x^{\deg f} f(1/x)$. Thus, assuming Conjecture 4, we were able to verify Conjecture 1 by computation for all $q \leq 100\,000$.

8.4 Relations with $(-1, 0, 1)$ polynomials

Related to the above discussion, we observed that in two special cases the non-reciprocal factors seem to have a very specific form in that they have coefficients $-1, 0, 1$. These cases are

(1) $a(2^{t(p)pm} - 1; x)$ for $m \geq 1$ and primes $p \geq 3$; see Proposition 6;

(2) $a(2^{p-1} + 1; x)$ for a prime $p \geq 3$ which has 2 as a primitive root; see Proposition 8.

In the second case we observed, in addition, that the nonzero coefficients are alternating between ± 1. As this lies outside the scope of the present paper, we did not pursue this further.

8.5 Another irreducibility criterion

We conclude this section by mentioning an irreducibility criterion of a different nature from those discussed earlier. Building on an interesting irreducibility result of A. Cohn, which had earlier been extended by Brillhart, Filaseta and Odlyzko (1981) and by Filaseta (1988), M. R. Murty [14] proved the following result.

Proposition 21 (Murty). *Let $b \geq 2$ and let p be a prime with b-adic expansion $p = a_n b^n + a_{n-1} b^{n-1} + \cdots + a_1 b + a_0$. Then the polynomial $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ is irreducible.*

Since Stern polynomials are $(0, 1)$-polynomials, we immediately get the following consequence.

Corollary 13. *If $a(n; b)$ is prime for some integer $b \geq 2$, then $a(n; x)$ is irreducible.*

By (7) we know that in most cases the degree of $a(n; x)$ is $(n - 1)/2$, or close to it. Thus, even for $b = 2$ the integers $a(n; b)$ will soon get very large as n grows. Therefore in most cases primality testing will not be competitive in comparison with irreducibility testing algorithms as implemented in computer algebra systems such as Maple or Mathematica.

Acknowledgement

We thank the referee for several helpful suggestions.

Research was supported in part by the Natural Sciences and Engineering Research Council of Canada.

References

[1] K. Dilcher, L. Ericksen: The polynomials of Mahler and roots of unity. *Amer. Math. Monthly.* (To appear).

[2] K. Dilcher, L. Ericksen: Identities and restricted quotients in the Stern sequence. (In preparation).

[3] K. Dilcher, K.B. Stolarsky: A polynomial analogue to the Stern sequence. *Int. J. Number Theory* 3 (1) (2007) 85–103.

[4] K. Dilcher, K.B. Stolarsky: Stern polynomials and double-limit continued fractions. *Acta Arith.* 140 (2) (2009) 119–134.

[5] M. Filaseta, K. Ford, S Konyagin: On an irreducibility theorem of A. Schinzel associated with coverings of the integers. *Illinois J. Math.* 44 (3) (2000) 633–643.

[6] M. Filaseta, M. Matthews (Jr.): On the irreducibility of 0,1-polynomials of the form $f(x)x^n + g(x)$. *Colloq. Math.* 99 (1) (2004) 1–5.

[7] C. Finch, L. Jones: On the irreducibility of $\{-1, 0, 1\}$-quadrinomials. *Integers* 6 (2006). A16, 4 pp.

[8] S. Klavžar, U. Milutinović, C. Petr: Stern polynomials. *Adv. in Appl. Math.* 39 (2007) 86–95.

[9] D.H. Lehmer: On Stern's diatomic series. *Amer. Math. Monthly* 36 (1929) 59–67.

[10] W. Ljunggren: On the irreducibility of certain trinomials and quadrinomials. *Math. Scand.* 8 (1960) 65–70.

[11] K. Mahler: On the zeros of a special sequence of polynomials. *Math. Comp.* 39 (159) (1982) 207–212.

[12] I. Mercer: Newman polynomials, reducibility, and roots on the unit circle. *Integers* 12 (2012) 503–519.

[13] W.H. Mills: The factorization of certain quadrinomials. *Math. Scand.* 57 (1985) 44–50.

[14] M.R. Murty: Prime numbers and irreducible polynomials. *Amer. Math. Monthly* 109 (2002) 452–458.

[15] OEIS Foundation Inc. (2011): The On-Line Encyclopedia of Integer Sequences. `http://oeis.org`

[16] T.J. Rivlin: *Chebyshev Polynomials,* second edition. Wiley, New York (1990).

[17] A. Schinzel: On the factors of Stern polynomials (remarks on the preceding paper of M. Ulas). *Publ. Math. Debrecen* 79 (1–2) (2011) 83–88.

[18] E.S. Selmer: On the irreducibility of certain trinomials. *Math. Scand.* 4 (1956) 287–302.

[19] B. Tuckerman: Factorization of $x^{2n} + x^n + 1$ using cyclotomic polynomials. *Math. Magazine* 42 (1) (1969) 41–42.

[20] M. Ulas: On certain arithmetic properties of Stern polynomials. *Publ. Math. Debrecen* 79 (1–2) (2011) 55–81.

[21] A.R. Vargas: Zeros and convergent subsequences of Stern polynomials. *J. Math. Anal. Appl.* 398 (2013) 630–637.

[22] E.W. Weisstein: Cyclotomic Polynomial. From MathWorld–A Wolfram Web Resource. `http://mathworld.wolfram.com/CyclotomicPolynomial.html`

Authors' addresses:
Karl Dilcher: Department of Mathematics and Statistics, Dalhousie University, Halifax, Nova Scotia, B3H 4R2, Canada

E-mail: `dilcher@mathstat.dal.ca`

Larry Ericksen: P.O. Box 172, Millville, NJ 08332-0172, USA

E-mail: `LE22@cornell.edu`

Received: 28 October, 2013
Accepted for publication: 9 May, 2014
Communicated by: Olga Rossi

Contents of Previous Volumes*

No. 1, Vol. 18 (2010)

No. 1, Vol. 19 (2011)

*Full texts in pdf are available free on-line at http://cm.osu.cz.

No. 2, Vol. 21 (2013)

Aims and Scope

Communications in Mathematics publishes original research papers and survey papers in pure and applied mathematics, preferably in the following areas: *algebraic structures*, *calculus of variations*, *combinatorics*, *control and optimization*, *cryptography*, *differential equations*, *differential geometry*, *fuzzy logic and fuzzy set theory*, *global analysis*, *history of mathematics*, *mathematical physics*, *number theory*.

To be acceptable for publication papers must make original and significant contributions to one or more of the topics within these areas. High quality review papers of interest to a wide range of scientists in mathematics and its applications are welcome.

Instructions for Authors

Articles submitted for publication must be written in English, and produced in TEX, preferably using Publisher's document style based upon LATEX.

Articles should contain the title, authors' names and addresses, full contact details of the corresponding author, the abstract, keywords, and 2010 MS classification numbers.

Footnotes should be avoided if possible, but if they are included they should be brief, and numbered consecutively.

References must be as complete as possible and put in alphabetical order.

Submission to the journal proceeds totally on-line through the web page `http://cm.osu.cz`, and the author is guided stepwise through the procedure. To complete the submission successfully, the author is required to choose one of the Editors to handle the paper. He/she should choose an Editor whose areas fit as closely as possible with the topic of the paper. All correspondence, including notification of the Editor's decision, requests for revision, and proofs takes place by e-mail.

Communications in Mathematics, Volume 22, 2014, is the continuation of the journal *Acta Mathematica et Informatica Universitatis Ostraviensis*, ISSN 1211-4774 (1993–2003), and of the journal *Acta Mathematica Universitatis Ostraviensis*, ISSN 1214-8148 (2004–2009).

Published by The University of Ostrava
June 2014

Typeset by TEX

ISSN 1804-1388 (Print), ISSN 2336-1298 (Online)

Contents

Research papers

Available online at
http://cm.osu.cz