ADDITIVE COMBINATORICS, CUBIC FORMS ON $\mathbb{F}_2^n$ AND HURWITZ SQUARE IDENTITIES

SOPHIE MORIER-GENOUD AND VALENTIN OVSIENKO

Abstract. We apply the Hurwitz-Radon theory of square identities to additive combinatorics of sumsets in $\mathbb{F}_2^n$. We fix an arbitrary cubic function α on $\mathbb{F}_2^n$ and obtain information about the size and structure of a set $A \subset \mathbb{F}_2^n$ satisfying $\alpha\big|_{(A+A)\setminus\{0\}} \equiv 1$. We then consider two differrent sets, A and B, and obtain a lower bound for the size of the sumset $A+B$ under a similar condition for the additive quadruples. Our main tool is the non-associative algebra associated to a cubic function on $\mathbb{F}_2^n$.

1. Introduction

In this paper, we investigate relationship between two different subjects: additive combinatorics of sumsets in $\mathbb{F}_2^n$ and the theory of "square identities" (or "composition of quadratic forms"), initiated by Hurwitz [14]. Starting from two sets, $A, B \subset \mathbb{F}_2^n$, and assuming some restriction on additive quadruples in $A \times A \times B \times B$, we construct a square identity. The main ingredient of our construction is a cubic binary (or boolean) form. This construction allows us to establish several statements about sizes of sumsets, as well as statements about binary cubics.

We will need to fix the notation and recall several standard notions of combinatorics over $\mathbb{F}_2$.

- An element $x \in \mathbb{F}_2^n$ is represented as an n-tuple of 0 and 1: $x = (x_1, \dots, x_n)$. The *Hamming weight* $\mathrm{wt}(x)$ is the number of non-zero components x_i.
- A *cubic form* on $\mathbb{F}_2^n$ is a function $\alpha : \mathbb{F}_2^n \to \mathbb{F}_2$ of the form
$$\alpha(x) = \sum_{1 \le i \le j \le k \le n} \alpha_{ijk}\, x_i x_j x_k,$$
where the coefficients $\alpha_{ijk} \in \{0, 1\}$. Note that, over $\mathbb{F}_2$, we have $x_i^2 = x_i$ and therefore every cubic form can be viewed as a *homogeneous* form.
- The cardinality of a set $A \subset \mathbb{F}_2^n$ is denoted by $|A|$.
- Given two sets $A, B \subset \mathbb{F}_2^n$, the *sumset* $A+B$ is the set of elements of the form $a+b$ with $a \in A, b \in B$.

We also recall two central notions of the Hurwitz theory.

- A square identity of size $[r, s, N]$ is an identity of the form
$$(a_1^2 + \dots + a_r^2)\,(b_1^2 + \dots + b_s^2) = c_1^2 + \dots + c_N^2,$$
where c_i are bilinear expressions in a_j and b_k with coefficients in $\mathbb{Z}$. In [14], Hurwitz formulated his famous problem to determine all the triples (r, s, N) such that there exists an identity of size $[r, s, N]$. The values (r, s, N) are *optimal* if r and s cannot be increased and N cannot be decreased.

- The *Hurwitz-Radon function* ρ is a function on the set of natural numbers $\rho : \mathbb{N} \to \mathbb{N}$. If $N = 2^n(2m+1)$, then $\rho(N) = \rho(2^n)$ (i.e., it depends only on the dyadic part of N), and the latter number is given by

$$\rho(2^n) = \begin{cases} 2n+1, & n \equiv 0 \mod 4 \\ 2n, & n \equiv 1,2 \mod 4 \\ 2n+2, & n \equiv 3 \mod 4. \end{cases}$$

- The celebrated Hurwitz-Radon theorem [15, 19]; see also [21], is formulated as follows: *there exists an identity of size* $[r, N, N]$ *if and only if* $r \leq \rho(N)$. This is the only case where the Hurwitz problem is solved. Importance of the Hurwitz problem is due to various applications of square identities and their direct relations to many areas of number theory, algebra, geometry and topology.

Starting from two sets A, B and their sumset $A+B$, and assuming some conditions in terms of (an arbitrary) cubic form α, we produce a square identity of size $[|A|, |B|, |A+B|]$. This approach was developed in [17, 16, 18] in order to construct new square identities. In this paper, we apply the same technique to, conversely, obtain information about additive combinatorics of sumsets from known results on square identities.

In Section 2, we formulate our results about sumsets. Sections 3-5 contain all our constructions: the correspondence between cubic functions on $\mathbb{F}_2^n$ and real non-associative algebras, the correspondence between sumsets and square identities, explicit examples of maximal subsets for Hurwitz-Radon identities. Section 6 completes the proofs.

2. The main results

2.1. **An upper bound.** The classical Freiman theorem [7] and its analog over $\mathbb{F}_2$ due to Ruzsa, as well as various generalizations (see [20, 6, 9] and [8] for a survey of the whole subject) provide information about the structure of a set A under a restriction on the cardinality of $A+A$. Our first result gives information about the cardinality of A under a restriction on the structure of $A+A$.

Theorem 1. *Given a cubic function* $\alpha : \mathbb{F}_2^n \to \mathbb{F}_2$ *and a set* $A \subset \mathbb{F}_2^n$*, if for every* $x, x' \in A$*, such that* $x \neq x'$*, one has* $\alpha(x+x') = 1$*, then*

$$|A| \leq \rho(2^n).$$

This bound is sharp, at least in the cases $n \equiv 1, 2$ *or* $3 \pmod 4$.

We will give examples of a cubic form α and sets A satisfying the condition $\alpha(x+x') = 1$ and the equality $|A| = \rho(2^n)$ (except for the case $n \equiv 0 \pmod 4$, where we have no examples).

An immediate combinatorial consequence of the above theorem is the following statement.

Corollary 2.1. *If for every* $x, x' \in A$*, such that* $x \neq x'$*, the weight* $\mathrm{wt}(x+x')$ *is not a multiple of* 4*, then* $|A| \leq \rho(2^n)$.

Note that replacing the condition multiple of 4 by another integer, say 3 or 5, one obtains that the maximal cardinality $|A|$ is at least quadratic in n. The value 4 is the only value for which the upper bound is linear in n.

The following statement is a refinement of Theorem 1.

Corollary 2.2. *If for every $x \neq x' \in A$ one has $\alpha(x+x') = 1$, and $A \subset V$, where V is an affine subspace of $\mathbb{F}_2^n$, then $|A| \leq \rho(|V|)$.*

As above, we can replace the assumption $\alpha(x+x') = 1$ by a more restrictive one that $\mathrm{wt}(x+x')$ is not a multiple of 4.

Let us also emphasize that Theorem 1 implies that the set A cannot have a small doubling, namely $|A+A| \geq c|A|^2$, where c is a constant (asymptotically $\frac{1}{4}$).

2.2. **Additive quadruples.** Our second statement concerns so-called *additive quadruples*. If $A, B \subset \mathbb{F}_2^n$, four elements $x, x' \in A$, $y, y' \in B$ form an additive quadruple (x, x', y, y') if

$$x + x' + y + y' = 0.$$

We call an additive quadruple *proper* if $x \neq x'$ and $y \neq y'$.

Theorem 2. *Let $A, B \subset \mathbb{F}_2^n$ with $|A| \leq |B|$. If every proper additive quadruple (x, x', y, y') satisfies $\alpha(x+x') = 1$, then*

$$|A+B| \geq \Omega(|A|^{\frac{6}{5}}).$$

As above, the condition $\alpha(x+x') = 1$ can be replaced by the condition that the weight $\mathrm{wt}(x+x')$ is not a multiple of 4.

The Balog-Szemerédi-Gowers theorem [3, 11], in the $\mathbb{F}_2^n$ case (see [9]) states, roughly speaking, that the sumset $A+B$ grows slowly, provided there are "many" additive quadruples (of order $|A|^3$). The above result is a sort of converse statement.

2.3. **A few properties of cubic forms.** Classification of boolean cubic forms on $\mathbb{F}_2^n$, modulo the action of the linear group $\mathrm{GL}(m, 2)$ is unknown for $n > 9$; see [12, 4] and the website http://langevin.univ-tln.fr/project/. The above theorems provide invariants of cubic forms, which are of course not sufficient for their classification. Given a cubic form α, the maximal cardinality of a set $A \subset \mathbb{F}_2^n$ such that $\alpha(x+x') = 1$ for all $x \neq x' \in A$, is obviously an invariant of α. The maximal cardinality of $|A+B|$, where A, B are as in Theorem 2, is also an invariant.

We also prove the following simple property of cubic binary forms.

Proposition 2.3. *If $V \subset \mathbb{F}_2^n$ is an (affine) subspace such that $\alpha(v) = 1$ for all non-zero $v \in V$, then $\dim(V) \leq 3$.*

2.4. **The method.** Hurwitz's problem remains widely open and no explicit formulas or asymptotic for the optimal triplets (r, s, N) are known in general. However, the problem is old and much information is available; see [21].

Our method is based on the algebraic constructions developed in [17, 16, 18]. For every cubic form α on $\mathbb{F}_2^n$, we construct a real non-associative algebra with basis $\{e_x \,|\, x \in \mathbb{F}_2^n\}$ and the product encoded by α. The square identities are realized in the form $||a||\,||b|| = ||ab||$, provided a and b are chosen in "good" subspaces of the algebra, and where $||.||$ is the Euclidean norm.

The following statement already proved in [16] will be explained in the sequel.

Theorem 3. *Given subsets $A, B \subset \mathbb{F}_2^n$ and a cubic form α, if for every proper additive quadruple (x, x', y, y') one has $\alpha(x+x') = 1$, then there exists a square identity of size given by the cardinalities: $[|A|, |B|, |A+B|]$.*

Note that this statement is a refinement of Yuzvinsky's theorem [23] (see also [21], Theorem 13.A.1, p. 286). Yuzvinsky did not consider cubic forms and his result gives a lower bound for $|A+B|$ in terms of the Hopf-Stiefel function.

3. Cubic functions and twisted group algebras

In this section, we develop our technique and establish the relationship between cubic forms and non-associative algebras.

3.1. **Examples of cubic functions on $\mathbb{F}_2^n$.** The space of (boolean) functions on $\mathbb{F}_2^n$ with values in $\mathbb{F}_2$ is isomorphic to the quotient space of the space of polynomials with coefficients in $\mathbb{F}_2$ by the ideal generated by the relations $x_i^2 = x_i$, namely

$$\mathbb{F}_2\left[x_1, \ldots, x_n\right] / \left(x_i^2 - x_i \,:\, i = 1, \ldots, n\right).$$

Every boolean function can be expressed as a polynomial in variables $(x_1, \ldots, x_n)$ with coefficients in $\mathbb{F}_2$, but not in a unique way. We will be interested in polynomials of degree ≤ 3.

Example 3.1. Let us introduce a very special cubic function $\alpha_{\mathbb{O}}$ on $\mathbb{F}_2^n$: it is equal to 1 everywhere except for the vectors of weight $\mathrm{w}t(x)$ which is proportional to 4:

$$\alpha_{\mathbb{O}}(x) = \begin{cases} 0, & \text{if } \mathrm{w}t(x) \equiv 0 \pmod 4 \\ 1, & \text{otherwise.} \end{cases}$$

The explicit coordinate formula can be written as follows:

$$\alpha_{\mathbb{O}}(x) = \sum_{1 \leq i < j < k \leq n} x_i x_j x_k + \sum_{1 \leq i < j \leq n} x_i x_j + \sum_{1 \leq i \leq n} x_i, \tag{3.1}$$

that is, $\alpha_{\mathbb{O}}$ is just the total sum of the monomials of degree $1, 2$ and 3. Since $x_i = x_i^2 = x_i^3$, this function can be viewed as a homogeneous cubic form.

The form $\alpha_{\mathbb{O}}$ is special since it is invariant with respect to the action of the group $\mathfrak{S}_n$ of permutations of the coordinates. Functions of this type are sometimes called *counting functions*.

Example 3.2. Another $\mathfrak{S}_n$-invariant cubic form is

$$\alpha_0(x) = \sum_{1 \leq i < j < k \leq n} x_i x_j x_k$$

that vanishes for all x except for $\mathrm{w}t(x) = 4m + 1$.

The quadratic form

$$q(x) = \sum_{1 \leq i < j \leq n} x_i x_j + \sum_{1 \leq i \leq n} x_i$$

is equal to 1 if $\mathrm{w}t(x) = 4m + 1$ or $4m + 2$.

Let us give the comparative table of values of the above forms.

(3.2)

$\mathrm{w}t(x)$	1	2	3	4	5	6	7	8	$\cdots$
q	1	1	0	0	1	0	0	0	$\cdots$
$\alpha_{\mathbb{O}}$	1	1	1	0	1	1	1	0	$\cdots$
α_0	1	0	0	0	1	0	0	0	$\cdots$

The above forms, especially $\alpha_{\mathbb{O}}$ will be useful to construct non-associative algebras that play essential role for this work.

3.2. **Twisted group algebras.** We now recall the definition of a twisted group algebra over an abelian group. We restrict ourselves to the case of the group $\mathbb{F}_2^n$, and the ground field is $\mathbb{R}$; we refer to [5] for the general theory.

Definition 3.3. Let $f : \mathbb{F}_2^n \times \mathbb{F}_2^n \to \mathbb{F}_2$ be an arbitrary function of two variables. The *twisted group algebra* over $\mathbb{F}_2^n$ associated to f is the real 2^n-dimensional algebra with basis $\{e_x \,|\, x \in \mathbb{F}_2^n\}$ and the product given by

$$e_x \cdot e_{x'} = (-1)^{f(x,x')}\, e_{x+x'}.$$

This algebra is denoted by $(\mathbb{R}[\mathbb{F}_2^n], f)$.

This algebra is, in general, neither commutative nor associative. The non-commutativity is measured by the function

$$\beta(x, y) := f(x, y) + f(y, x),$$

while the non-associativity is measured by the function

$$\delta f(x, y, z) := f(y, z) + f(x + y, z) + f(x, y + z) + f(x, y).$$

Many classical algebras, such as the algebras of quaternions $\mathbb{H}$, of octonions $\mathbb{O}$, and, more generally, the Clifford algebras and the Cayley-Dickson algebras, can be realized as twisted group algebras over $\mathbb{F}_2^n$; see [1, 2].

The associativity condition, $\delta f = 0$, is too restrictive and does not lead to algebras interesting for combinatorics. For instance, the Clifford algebras are too simple for our purpose. On the other hand, algebras with $\delta f \neq 0$ can be very hard to handle, this is the case of the Cayley-Dickson algebras. In [17] we introduced an intermediate condition that is the symmetrized function of f is a 2-cocycle, i.e.,

$$\delta\beta = 0.$$

This leads to series of new algebras. We have proved in [17] that, if β as above is a 2-cocycle, then it is necessarily a coboundary, i.e., $\beta = \delta\alpha$, and α has to be a cubic form on $\mathbb{F}_2^n$. Moreover, α determines the twisted group algebra up to isomorphism.

We summarize the approach in the following short way.

3.3. **From cubic forms to algebras.** There is a canonical way to construct a twisted group algebra out of an arbitrary cubic form α.

Proposition 3.4. *Given a cubic form α, there exists a function f satisfying the conditions:*

(a) *First polarization formula:*

$$f(x, y) + f(y, x) = \alpha(x + y) + \alpha(x) + \alpha(y).$$

(b) *Second polarization formula:*

$$f(x, y) + f(x, y + z) + f(x + y, z) + f(y, z) = \\ \alpha(x + y + z) + \alpha(x + y) + \alpha(x + z) + \alpha(y + z) + \alpha(x) + \alpha(y) + \alpha(z).$$

(c) *Linearity of f in 2nd variable:*

$$f(x, y + y') = f(x, y) + f(x, y').$$

(d) *Reconstruction of α from f:*

$$f(x, x) = \alpha(x).$$

Proof. The existence of f follows from an explicit formula. We replace every monomial in α according to the following rule:

$$\begin{aligned} x_ix_jx_k &\longmapsto x_ix_jy_k + x_iy_jx_k + y_ix_jx_k, \\ x_ix_j &\longmapsto x_iy_j, \\ x_i &\longmapsto x_iy_i. \end{aligned} \tag{3.3}$$

where $i < j < k$, and obtain this way a function f in two arguments, satisfying the properties (a)-(d). □

Example 3.5. Our main examples are related to the octonions and Clifford algebras.

(1) The function $f_{\mathbb{O}} : \mathbb{F}_2^n \times \mathbb{F}_2^n \to \mathbb{F}_2$ corresponding to the form $\alpha_{\mathbb{O}}$ is as follows.

$$f_{\mathbb{O}}(x,y) = \sum_{1\le i<j<k\le n} (x_ix_jy_k + x_iy_jx_k + y_ix_jx_k) + \sum_{1\le i\le j\le n} x_iy_j$$

for all $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$, elements of $\mathbb{F}_2^n$. When $n = 3$, the corresponding twisted group algebra is isomorphic to the classical algebra $\mathbb{O}$ of octonions.

(2) The twisted group algebra $(\mathbb{R}[\mathbb{F}_2^n], f_q)$, where f_q is the function corresponding to the quadratic form q from Example 3.2, is isomorphic to the Clifford algebra $\mathrm{Cl}_{0,n}$.

The constructed algebras $(\mathbb{R}[\mathbb{F}_2^n], f_{\mathbb{O}})$ generalize the algebra of octonions (and the Clifford algebras). These algebras were introduced and studied in [17]. Their properties are completely different from those of Cayley-Dickson algebras.

Remark 3.6. Given a cubic form α, the choice of a function f is unique modulo a coboundary. More precisely, two functions of two variables, f and f', correspond to the same cubic form α, if and only if $f + f'$ is a coboundary. The corresponding twisted group algebras are isomorphic.

3.4. **Why cubic functions?** One cannot choose a polynomial of degree ≥ 4, instead of a cubic function, in order to construct a twisting function f satisfying properties (a)-(d) of Proposition 3.4. Indeed, let us apply the differential δ to the equation in property (b). Since $\delta^2 = 0$, one obtains after a short computation:

$$\begin{aligned} 0 = \ & \alpha(x+y+z+t) \\ & +\alpha(x+y+z)+\alpha(x+y+t)+\alpha(x+z+t)+\alpha(y+z+t) \\ & +\alpha(x+y)+\alpha(x+z)+\alpha(x+t)+\alpha(y+z)+\alpha(y+t)+\alpha(z+t) \\ & +\alpha(x)+\alpha(y)+\alpha(z)+\alpha(t). \end{aligned}$$

This is exactly the condition that α is a polynomial of degree at most 3.

3.5. **A criterion of existence of α.** An equivalent way to express the condition $\delta\beta = 0$ (and therefore the existence of α) is to require that the ternary function $\delta f(x,y,z)$ is symmetric in its arguments x, y, z. This means that the non-associativity is not "too wild". In particular, the algebras are alternative.

4. From algebra to square identities

4.1. **The Euclidean norm.** Consider a twisted group algebra $(\mathbb{R}[\mathbb{F}_2^n], f)$. Our plan is to define the Euclidean norm (or absolute value) on the space $\mathbb{R}[\mathbb{F}_2^n] \cong \mathbb{R}^{2^n}$ and investigate its compatibility with the algebra structure.

Every element of the algebra is a linear combination of the basis elements

$$a = \sum_{x \in \mathbb{F}_2^n} a_x \, e_x,$$

with (arbitrary) real coefficients $a_x \in \mathbb{R}$. We set

$$||a||^2 := \sum_{x \in \mathbb{F}_2^n} a_x^2.$$

Consider two sets $A, B \subset \mathbb{F}_2^n$ and the coordinate subspaces $\mathcal{A}$ and $\mathcal{B} \subset (\mathbb{R}[\mathbb{F}_2^n], f)$:

$$\big\{a \mid a = \sum_{x \in A} a_x \, e_x\big\} \quad \text{and} \quad \big\{b \mid b = \sum_{y \in B} b_y \, e_y\big\}.$$

Our next task is to determine a necessary and sufficient condition on A and B that guarantees:

$$||a||^2 \, ||b||^2 = ||a\, b||^2, \tag{4.1}$$

for all $a \in \mathcal{A}$ and $b \in \mathcal{B}$. Observe that the equation (4.1) is nothing but a square identity of size $[|A|, |B|, |A+B|]$.

4.2. **The normed subspaces.** This idea goes back to Yuzvinsky [23], it was also used in [17, 16, 18]. The following statement was obtained in all these references, we give here a proof for the sake of consistency.

Lemma 4.1. *The condition (4.1) is satisfied if and only if for all $x \neq x' \in A$ and $y \neq y' \in B$ such that $x + x' + y + y' = 0$, one has:*

$$f(x,y) + f(x,y') + f(x',y) + f(x',y') = 1. \tag{4.2}$$

Proof. The product of the norm in the left-hand-side of (4.1) is:

$$||a||^2 \, ||b||^2 = \sum_{x,y} a_x^2 \, b_y^2.$$

On the other hand, the product of two elements is given by

$$a\, b = \sum_{x+y} (-1)^{f(x,y)} \, a_x \, b_y \, e_{x+y} = \sum_z \left(\sum_{x+y=z} (-1)^{f(x,y)} \, a_x \, b_y \right) e_z.$$

The Euclidean norm of this element is

$$||a\, b||^2 = \sum_{x,y} a_x^2 \, b_y^2 + \sum_{\substack{x \,+\, y \,=\, x' \,+\, y' \\ x \,\neq\, x'}} (-1)^{f(x,y)+f(x',y')} \, a_x \, b_y \, a_{x'} \, b_{y'}.$$

The monomial $a_x \, b_y \, a_{x'} \, b_{y'}$ in the second summand appears twice, and has total coefficient

$$(-1)^{f(x,y)+f(x',y')} + (-1)^{f(x,y')+f(x',y)}.$$

This coefficient vanishes if and only if the condition (4.2) holds. □

4.3. **The norm condition in terms of the cubic.** Suppose now that the algebra $(\mathbb{R}[\mathbb{F}_2^n], f)$ was constructed out of a cubic form α; see Section 3.3. The equation (4.2) then drastically simplifies.

Lemma 4.2. *The equation (4.2) reads:* $\alpha(x+x')=1$.

Proof. Use linearity of f in second argument and substitute $y'=x+x'+y$ to the left-hand-side of (4.2). After cancellation one has

$$f(x,x)+f(x,x')+f(x',x)+f(x',x')=\alpha(x)+\beta(x,x')+\alpha(x')=\alpha(x+x'),$$

thanks to the properties (a) and (d) of Proposition 3.4. □

5. Construction of Hurwitzian sets

In this section, we fix the cubic form $\alpha_{\mathbb{O}}$ given by (3.1) and study the sets $A\subset\mathbb{F}_2^n$ satisfying the condition $\alpha_{\mathbb{O}}(x+x')=1$ for all distinct $x,x'\in A$. In other words, the weight $\mathrm{w}t(x+x')$ is not a multiple of 4. We are interested in the sets A of cardinality $|A|=\rho(2^n)$. Such sets were already considered in [16] where they were called *Hurwitzian sets*. In this section, will give explicit constructions of Hurwitzian sets. In particular, we discuss a relation to the binary Hadamard matrices.

5.1. **Cases** $n\equiv 1,2\,(\mathrm{mod}\ 4)$**.** In this case, $\rho(2^n)=2n$. The following choice of a Hurwitzian set is perhaps the most obvious. Choose the following set:

$$A=\{0,\ e_1,\ e_2,\ \dots,\ e_n,\ e_1+e_2,\ e_1+e_3,\ \dots,\ e_1+e_n\}.$$

For all $x,x'\in A$, the weight of the sum satisfies $\mathrm{w}t(x+x')\le 3$, and thus $\alpha_{\mathbb{O}}(x+x')=1$, provided $x+x'\neq 0$. Therefore A is a Hurwitzian set.

Note that the above choice is not unique. However, it is easy to see that the set A is the only Hurwitzian set which is a "shift-minimal downset" according to the terminology of [10].

5.2. **Case** $n\equiv 3\,(\mathrm{mod}\ 4)$**.** In this case, $\rho(2^n)=2n+2$ which is the most interesting situation for many reasons.

Consider the element of maximal weight:

$$\omega=(11\dots1)=e_1+\dots+e_n.$$

One can choose the above set A, completed by ω and $e_1+\omega$. Let us give a more symmetric example.

Choose the set

$$A=\{0,\ \omega,\ e_1,\ e_2,\ \dots,\ e_n,\ e_1+\omega,\ e_2+\omega,\ \dots,\ e_n+\omega\}.$$

The weight of a non-zero element of the sumset $A+A$ can be one of the following four values: $1,2,n-1$, or $n-2$. Since this is never a multiple of 4, we conclude that A is a Hurwitzian set. Moreover, it is not difficult to show that the above set is the only Hurwitzian set invariant with respect to the group of permutations $\mathfrak{S}_n$.

5.3. **Another choice in the case $n \equiv 3 \pmod 8$, relation to the Hadamard matrices.** The case $n \equiv 3 \pmod 8$ is a subcase of the above one. Remarkable, there is a choice of Hurwitzian set based on the classical Hadamard matrices.

Recall that a *Hadamard matrix* is an $m \times m$-matrix H with entries ± 1, such that ${}^tHH = m\mathcal{I}$, where ${}^tH$ is the transpose of H and $\mathcal{I}$ is the identity matrix. It is known that a Hadamard matrix can exist only if $m = 1, 2$ or $m = 4s$; existence for arbitrary s is the classical Hadamard conjecture.

The construction is as follows. We remove the first column and than consider two 11×12-matrices, H_1, H_2 with entries $0, 1$. The matrix H_1 is obtained by replacing 1 by 0 and -1 by 1, the matrix H_2 is obtained by replacing -1 by 0.

Lemma 5.1. *The rows of H_1 and H_2 form a Hurwitzian set in $\mathbb{F}_2^{4s-1}$, provided s is odd.*

Proof. It follows from the definition of a Hadamard matrix that every sum of two distinct rows of H_1 is of weight $2s$, and similarly for H_2. The sum of a row of H_1 with a row of H_2 is of weight $2s - 1$ or $4s - 1$. □

Example 5.2. The (unique up to equivalence) 12×12 Hadamard matrix H corresponds to the following 12×11 binary matrices:

$$H_1 = \begin{pmatrix} 1&1&1&1&1&1&1&1&1&1&1\\ 0&1&0&1&1&1&0&0&0&1&0\\ 0&0&1&0&1&1&1&0&0&0&1\\ 1&0&0&1&0&1&1&1&0&0&0\\ 0&1&0&0&1&0&1&1&1&0&0\\ 0&0&1&0&0&1&0&1&1&1&0\\ 0&0&0&1&0&0&1&0&1&1&1\\ 1&0&0&0&1&0&0&1&0&1&1\\ 1&1&0&0&0&1&0&0&1&0&1\\ 1&1&1&0&0&0&1&0&0&1&0\\ 0&1&1&1&0&0&0&1&0&0&1\\ 1&0&1&1&1&0&0&0&1&0&0 \end{pmatrix} \quad H_2 = \begin{pmatrix} 0&0&0&0&0&0&0&0&0&0&0\\ 1&0&1&0&0&0&1&1&1&0&1\\ 1&1&0&1&0&0&0&1&1&1&0\\ 0&1&1&0&1&0&0&0&1&1&1\\ 1&0&1&1&0&1&0&0&0&1&1\\ 1&1&0&1&1&0&1&0&0&0&1\\ 1&1&1&0&1&1&0&1&0&0&0\\ 0&1&1&1&0&1&1&0&1&0&0\\ 0&0&1&1&1&0&1&1&0&1&0\\ 0&0&0&1&1&1&0&1&1&0&1\\ 1&0&0&0&1&1&1&0&1&1&0\\ 0&1&0&0&0&1&1&1&0&1&1 \end{pmatrix}$$

(which are related to the extended Golay code). The rows of the matrices H_1 and H_2 constitute a Hurwitzian set in $\mathbb{F}_2^{11}$ of cardinality 24.

5.4. **Case $n \equiv 0 \pmod 4$.** Recall that $\rho(2^n) = 2n + 1$ in this case. We have no construction of Hurwitzian set. Moreover, we are convinced that a similar situation holds for any cubic form.

Conjecture 1. *Given a boolean cubic function α on $\mathbb{F}_2^n$ with $n \equiv 0 \pmod 4$, there is no set A such that $\alpha\big|_{(A+A)\setminus\{0\}} \equiv 1$ and $|A| = 2n + 1$.*

6. Proof of the main results

We are ready to complete the proof of the results formulated in Section 2.

6.1. **Proof of Theorem 1 and Corollaries 2.1 and 2.2.** Fix an arbitrary cubic form α, and let $A \subset \mathbb{F}_2^n$ be a set such that

$$\alpha \,\big|_{\,(A+A)\setminus\{0\}} \equiv 1.$$

Lemmas 4.1 and 4.2 then imply $||a||\,||b|| = ||ab||$ for all $a \in \mathcal{A}$ and arbitrary $b \in \mathbb{F}_2^n$. We therefore obtain a square identity of size $[|A|, 2^n, 2^n]$. The Hurwitz-Radon Theorem implies

that $|A| \leq \rho(2^n)$. This bound is sharp as follows from the constructions of Hurwitzian sets; see Section 5. Theorem 1 is proved.

Choosing $\alpha = \alpha_{\mathbb{O}}$ as in formula (3.1), one obtains Corollary 2.1.

Now we deduce Corollary 2.2. It will suffice to consider linear subspace V. Let $A \subset V$, the linear subspace $V \subset \mathbb{F}_2^n$ is itself isomorphic to $\mathbb{F}_2^m$ for some $m \leq n$. Given a cubic form $\alpha_{\mathbb{O}}$ on $\mathbb{F}_2^n$ and let α_V be its pull-back to V. One then has $\alpha_V(x+y) = 1$, for all $x, y \in A$ and concludes by the same arguments as above. Corollary 2.2 then follows.

6.2. **Proof of Theorem 2.** Fix, as above, an arbitrary cubic form α. Suppose that A and B are two subsets of same cardinality $|A| = |B| = r$, and such that for all proper additive quadruples (x, x', y, y'). By Theorem 3, one obtains an identity of size $[r, r, N]$, where $N = |A + B|$. The Hurwitz problem is still open in this particular case and even an asymptotic of the least value $N_{\min}$ as a function of r is not known exactly. However, it is known that asymptotically

$$C_1\, r^{\frac{6}{5}} \leq N_{\min}(r) \leq C_2\, \frac{r^2}{\log(r)}.$$

where C_1 and C_2 are some constants. The upper bound follows easily from the Hurwitz-Radon theorem, and the lower bound was recently obtained in [13]. This is precisely the statement of Theorem 2.

6.3. **Proof of Proposition 2.3.** If $V \subset \mathbb{F}_2^n$ is a subspace such that $\alpha(v) = 1$ for all non-zero $v \in V$, then there exists a square identity of size $[|V|, |V|, |V|]$. The famous theorem of Hurwitz [14] states that there is an identity of size $[N, N, N]$, if and only if $N = 1, 2, 4$ or 8. It follows that $\dim(V) \leq 3$.

Acknowledgments. This project was partially supported by the PICS05974 "PENTAFRIZ" of CNRS.

References

[1] H. Albuquerque, S. Majid, *Quasialgebra structure of the octonions*, J. Algebra **220** (1999), 188–224.
[2] H. Albuquerque, S. Majid, *Clifford algebras obtained by twisting of group algebras*, J. Pure Appl. Algebra **171** (2002), 133–148.
[3] A. Balog and E. Szemerédi, *A statistical theorem of set addition*, Combinatorica **14** (1994), 263–268.
[4] E. Brier, P. Langevin, *The classification of boolean cubics of nine variables*, 2003 IEEE Information Theory Workshop, La Sorbonne, Paris, France (2003).
[5] S.B. Conlon, *Twisted group algebras and their representations*, J. Austr. Math. Soc. **4** (1964), 152–173.
[6] J.-M. Deshouillers, F. Hennecart, A. Plagne, *On small sumsets in* $(\mathbb{Z}/2\mathbb{Z})^n$, Combinatorica **24** (2004), 53–68.
[7] G. Freiman, Foundations of a structural theory of set addition, Translations of Mathematical Monographs **37**, Amer. Math. Soc., Providence, 1973.
[8] B. J. Green, *Finite field models in arithmetic combinatorics*, in: Surveys in Combinatorics 2005, London Math. Soc. Lecture Notes **327**, 1–27.
[9] B. Green, T. Tao, *A note on the Freiman and Balog-Szemerédi-Gowers theorems in finite fields*, J. Aust. Math. Soc. **86** (2009), 61–74.
[10] B. Green, T. Tao, *Freiman's theorem in finite fields via extremal set theory,* Combin. Probab. Comput. **18** (2009), 335–355.
[11] W. T. Gowers, A new proof of Szemerédi's theorem for arithmetic progressions of length four, Geom. Funct. Anal. **8** (1998), 529–551.
[12] X.-D. Hou, $\mathrm{GL}(m, 2)$ *acting on* $R(r, m)/R(r-1, m)$, Discrete Math. **149** (1996), 99–122.
[13] P. Hrubes, A. Wigderson, A. Yehudayoff, *Non-commutative circuits and the sum-of-squares problem*, J. Amer. Math. Soc. **24** (2011), 871–898.

[14] A. Hurwitz, *Über die Komposition der quadratischen Formen von beliebig vielen Variablen*, Nahr. Ges. Wiss. Göttingen (1898), 309–316.
[15] A. Hurwitz, *Über die Komposition der quadratischen Formen*, Math. Ann. 88 (1922), 1–25.
[16] A. Lenzhen, S. Morier-Genoud, V. Ovsienko, *New solutions to the Hurwitz problem on square identities*, J. Pure Appl. Algebra **215** (2011), 2903–2911.
[17] S. Morier-Genoud, V. Ovsienko, *A series of algebras generalizing the octonions and Hurwitz-Radon identity*, Comm. Math. Phys. **306** (2011), 83–118.
[18] S. Morier-Genoud, V. Ovsienko, *Orthogonal designs and a cubic binary function*, IEEE Trans. Information Theory, **59**:3 (2013) 1583–1589.
[19] J. Radon, *Lineare Scharen orthogonaler Matrizen*, Abh. Math. Sem. Univ. Hamburg **1** (1922) 1–14.
[20] I. Z. Ruzsa, *An analog of Freiman's theorem in groups*, in: Structure theory of set addition, Astérisque No. 258 (1999), 323–326.
[21] D. Shapiro, Compositions of quadratic forms, Walter de Gruyter & Co., Berlin, 2000.
[22] T. Tao, V.H. Vu, Additive combinatorics. Cambridge Studies in Advanced Math., **105**, Cambridge University Press, Cambridge, 2006.
[23] S. Yuzvinsky, *Orthogonal pairings of Euclidean spaces*, Michigan Math. J. **28** (1981), 131–145.

Sophie Morier-Genoud, Université Paris 6, IMJ, Equipe Analyse Algébrique, UFR 929 de Mathématiques, 175 rue du Chevaleret, 75013 Paris, France

Valentin Ovsienko, CNRS, Institut Camille Jordan, Université Claude Bernard Lyon 1, 43 boulevard du 11 novembre 1918, 69622 Villeurbanne cedex, France

E-mail address: sophiemg@math.jussieu.fr, ovsienko@math.univ-lyon1.fr