J. reine angew. Math. **545** (2002), 67—95

Journal für die reine und angewandte Mathematik

Explicit constructions in the classical mean squares problem in irregularities of point distribution

Dedicated to the memory of Paul Erdős

By *W. W. L. Chen* at Sydney[1] and *M. M. Skriganov* at St Petersburg[2]

Abstract. We explicitly construct distributions $\mathscr{D}_N$ of N points in the n-dimensional unit cube U^n with the minimal order of the L_2-discrepancy $\mathscr{L}_2[\mathscr{D}_N] < C_n(\log N)^{\frac{1}{2}(n-1)}$, where the constant C_n is independent of N. The constructions are based on ideas from coding theory. In particular, we use codes over finite fields $\mathbb{F}_p$ with large weights simultaneously in two different metrics—the well known Hamming metric and a new non-Hamming metric arising recently in coding theory.

1. Introduction

Suppose that $\mathscr{A}_N$ is a distribution of $N > 1$ points, not necessarily distinct, in the n-dimensional unit cube $U^n = [0,1)^n$. The L_2-discrepancy $\mathscr{L}_2[\mathscr{A}_N]$ is defined by

$$\mathscr{L}_2[\mathscr{A}_N] = \left(\int_{U^n} |\mathscr{L}[\mathscr{A}_N; Y]|^2 \, \mathrm{d}Y \right)^{1/2},$$

where for every $Y = (y_1, \ldots, y_n) \in U^n$, the local discrepancy $\mathscr{L}[\mathscr{A}_N; Y]$ is given by

$$\mathscr{L}[\mathscr{A}_N; Y] = \#(\mathscr{A}_N \cap B_Y) - N \operatorname{vol} B_Y. \tag{1.1}$$

Here $B_Y = [0, y_1) \times \cdots \times [0, y_n) \subseteq U^n$ is a rectangular box of volume $\operatorname{vol} B_Y = y_1 \ldots y_n$, while $\#(\mathscr{S})$ denotes the number of points of a set $\mathscr{S}$, counted with multiplicity.

[1] Research supported by grants from the Australian Research Council and Macquarie University.
[2] Research supported by the Russian Fund of Fundamental Research 99-01-00106 and INTAS 00-429.

In 1954, Roth [20] established a lower bound for the L_2-discrepancy. For an arbitrary distribution $\mathscr{A}_N$ of N points in the unit cube U^n, we have

$$\mathscr{L}_2[\mathscr{A}_N] > c_n(\log N)^{\frac{1}{2}(n-1)}, \tag{1.2}$$

with a positive constant c_n depending only on the dimension n. This lower bound (1.2) turns out to be best possible. More precisely, for every dimension n and every $N > 1$, there exist distributions $\mathscr{B}_N$ of N points in the unit cube U^n such that

$$\mathscr{L}_2[\mathscr{B}_N] < C_n(\log N)^{\frac{1}{2}(n-1)}, \tag{1.3}$$

with a positive constant C_n depending only on the dimension n.

The first constructions of distributions $\mathscr{B}_N$ satisfying the inequality (1.3) were given in dimensions $n = 2$ and $n = 3$ by Davenport [6] and Roth [21] respectively, and in arbitrary dimensions by Roth [22]. For the early history of this problem, we refer the reader to Beck and Chen [1]. Other distributions $\mathscr{B}_N$ satisfying the upper bound (1.3) were also constructed by Chen [3], [4], Frolov [10], Dobrovol'skiĭ [7] and Skriganov [24], [25].

Until recently, apart from Davenport's construction in 1956 for dimension $n = 2$, all known constructions of point sets $\mathscr{B}_N$ for dimensions $n > 2$ with minimal order of the L_2-discrepancy (1.3) involve probabilistic arguments and are therefore not explicit. Note that the arguments in [4] and [7] involve discrete variables and can be considered to be effective. However, any implementation of the ideas in these two papers will be excessively complicated.

In this paper, we give a complete solution to the explicit construction problem in arbitrary dimensions. More precisely, we prove the following result.

Theorem. *Let $p \geqq 2n^2$ be a prime. Then for every $N > 1$, a distribution $\mathscr{D}_N$ of N points in the unit cube U^n can be constructed explicitly to satisfy the inequality*

$$\mathscr{L}_2[\mathscr{D}_N] < 2^{n+1}p^{2n}(\log N + 2n + 1)^{\frac{1}{2}(n-1)}.$$

We remark here that recently Larcher and Pillichshammer [12] have independently studied the problem for dimension $n = 3$, with an upper bound of order of magnitude $(\log N)(\log\log N)^{1/2}$. Their approach uses the point sets constructed in 1982 by Faure [8], and involves a lengthy analysis of the binomial coefficients involved.

Here we shall consider distributions which possess the structure of vector spaces over finite fields with respect to a certain p-ary arithmetic operation. Such point sets are distributed very uniformly in the unit cube with respect to the supremum norm if the corresponding vector spaces have large weights relative to a special metric, as shown in a paper of the second author [26], where the theory of uniform distribution was studied in detail in the context of coding theory. In the present paper, we shall show that such point sets satisfy the inequality (1.3) if the corresponding vector spaces have large weights simultaneously in two special metrics. We shall briefly discuss these ideas from [26], adapted to our present context, in the last two sections of this paper.

Another crucial concept of our approach is the use of suitably generalized Walsh functions. The classical Walsh functions were first used in our earlier investigations in [5]. The generalized Walsh functions, like the classical ones, form an orthonormal basis of the space $L_2(U^n)$, and can be used to obtain very convenient expressions for the local discrepancy (1.1). At the same time, such functions are additive characters of vector spaces over finite fields. This allows us to give a very precise evaluation of the corresponding L_2-discrepancy.

A pleasant surprise in our present approach is that we do not need the concept of the Davenport reflection principle, a crucial idea in [5] to replace probabilistic arguments in earlier work. This has the benefit of substantially simplifying the argument. It appears that the point sets we consider here have some "self-averaging" property which is not present in other point sets used before.

It is hoped that the approach given in the present paper may have wider applications in the theory of uniform distribution.

The paper is organized as follows. In Section 2, we shall discuss point sets that have the structure of vector spaces over finite fields, as well as two special metrics central to our argument. We shall state a number of lemmas crucial to the proof of our result. This section contains many ideas first discussed in [26] in a slightly different form, and we shall leave some of the detailed justifications until Sections 7 and 8. In Section 3, we shall show how we may combine the various lemmas in Section 2 to deduce the Theorem. Sections 4–6 are devoted to the establishment of the crucial Lemma 2D. In Section 4, we recall necessary facts on the generalized Walsh functions. These will be used in Section 5 to obtain Fourier-Walsh expansions for characteristic functions of intervals and rectangular boxes. We then deduce Lemma 2D in Section 6.

Throughout, the letter p denotes a prime satisfying certain conditions depending only on the dimension n, which we shall specify later. For this prime,

$$\mathbb{F}_p = \{0, 1, \ldots, p-1\}$$

denotes the finite field of residues modulo p. For convenience, $\mathbb{N}$ denotes the set of all positive integers, $\mathbb{N}_0$ denotes the set of all non-negative integers, $\mathbb{Q}$ denotes the set of all rational numbers, and $\mathbb{C}$ denotes the set of all complex numbers. If $z \in \mathbb{C}$, then $\bar{z} \in \mathbb{C}$ denotes its complex conjugate. Finally, if $\mathscr{S}$ is a finite set, then $\#(\mathscr{S})$ denotes the number of elements of $\mathscr{S}$.

Acknowledgments. We thank our friends Ralph Alexander, József Beck, Gerhard Larcher, Jiří Matoušek, Andy Pollington, Klaus Roth and Henryk Woźniakowski for our many interesting and valuable discussions. Our joint interest in this problem started a number of years ago when the second author visited Macquarie University in 1994 and 1996, and the first author visited the Steklov Mathematical Institute in 1995. We thank our colleagues both in Sydney and in St Petersburg for their hospitality and support. The generosity and hospitality of Maurice Dodson and his group at the University of York in 1997, and of Anand Srivastav and his group at the University of Kiel in 1998, enable us to continue our investigation with the luxury of being able to sit in the same continent instead of communicating by electronic mail from opposite ends of the world. We record our indebtedness to them all.

2. Nets and linear distributions

For $a, m \in \mathbb{N}_0$ satisfying $0 \leqq m < p^a$, we shall consider elementary intervals of the type

$$\Delta_a^m = \big[mp^{-a}, (m+1)p^{-a}\big).$$

For vectors $A = (a_1, \dots, a_n)$ and $M = (m_1, \dots, m_n)$ in $\mathbb{N}_0^n$ satisfying $0 \leqq m_j < p^{a_j}$ for every $j = 1, \dots, n$, we shall consider elementary boxes of the type

$$\Delta_A^M = \Delta_{a_1}^{m_1} \times \cdots \times \Delta_{a_n}^{m_n}. \tag{2.1}$$

Clearly any such box has volume $\operatorname{vol} \Delta_A^M = p^{-a_1 - \cdots - a_n}$.

Definition. Suppose that $s, \delta \in \mathbb{N}_0$ satisfy $0 \leqq \delta \leqq s$. A set $D \subset U^n$ of p^s points is called an (n, s, δ)-net (in base p) with deficiency δ if every elementary box Δ_A^M of volume $p^{\delta - s}$ contains precisely p^δ points of D.

These sets were first constructed by Sobol [27] and Faure [8]. For a systematic treatment of nets, including nets in arbitrary integer bases, see [16] or the survey [17]. It is almost trivial to show that for such a set D, the inequality

$$\mathscr{L}[D; Y] = O(p^\delta s^{n-1}) = O\big(p^\delta (\log N)^{n-1}\big) \tag{2.2}$$

holds for every $Y \in U^n$, where $N = \#(D) = p^s$ and where the implied constants in (2.2) depend only on n and p. It follows that such nets fill out the unit cube very uniformly as $s \to \infty$. For a proof of the inequality (2.2), see Section 3.2 of [1] or Section 3 of [16].

We shall also be concerned with a class of sets $D \subset U^n$ which possess the structure of vector spaces over the finite field $\mathbb{F}_p$.

For any $s \in \mathbb{N}_0$, let

$$\mathbb{Q}(p^s) = \{mp^{-s} \colon m = 0, 1, \dots, p^s - 1\}$$

and

$$\mathbb{Q}^n(p^s) = \{X = (x_1, \dots, x_n) \colon x_1, \dots, x_n \in \mathbb{Q}(p^s)\}.$$

Furthermore, we write

$$\mathbb{Q}(p^\infty) = \bigcup_{s=0}^{\infty} \mathbb{Q}(p^s) \quad \text{and} \quad \mathbb{Q}^n(p^\infty) = \bigcup_{s=0}^{\infty} \mathbb{Q}^n(p^s).$$

Points in $\mathbb{Q}^n(p^\infty)$ are sometimes called the p-ary rational points.

Observe that any $x \in [0, 1)$ can be represented in the form

$$x = \sum_{i=1}^{\infty} \eta_i(x) p^{-i}, \tag{2.3}$$

where the coefficients $\eta_i(x) \in \mathbb{F}_p = \{0, 1, \dots, p-1\}$ for every $i \in \mathbb{N}$. This representation is unique if we agree that the series in (2.3) is finite for every $x \in \mathbb{Q}(p^\infty)$.

For any two vectors $X = (x_1, \dots, x_n)$ and $Y = (y_1, \dots, y_n)$ in $\mathbb{Q}^n(p^\infty)$ and any two scalars $\alpha, \beta \in \mathbb{F}_p$, we write

$$\alpha X \oplus \beta Y = (\alpha x_1 \oplus \beta y_1, \dots, \alpha x_n \oplus \beta y_n) \in \mathbb{Q}^n(p^\infty) \tag{2.4}$$

by setting

$$\eta_i(\alpha x_j \oplus \beta y_j) = \alpha \eta_i(x_j) + \beta \eta_i(y_j) \pmod p$$

for every $i \in \mathbb{N}$ and $j = 1, \dots, n$. It is easy to see that with respect to the arithmetic operations (2.4), each set $\mathbb{Q}^n(p^s)$ forms a vector space of dimension ns over $\mathbb{F}_p$, while the set $\mathbb{Q}^n(p^\infty)$ forms an infinite dimensional vector space over $\mathbb{F}_p$.

Definition. We say that a subset $D \subseteqq \mathbb{Q}^n(p^\infty)$ is a linear distribution (in base p) if D is a subspace of the vector space $\mathbb{Q}^n(p^\infty)$.

In this paper, we shall only consider linear distributions of finitely many points. Clearly any such finite linear distribution is a subspace of the vector space $\mathbb{Q}^n(p^s)$ for all sufficiently large values of $s \in \mathbb{N}_0$.

Suppose now that $s \in \mathbb{N}_0$ is chosen and fixed. Then any $x \in \mathbb{Q}(p^s)$ can be represented also in the form

$$x = \sum_{i=1}^{s} \xi_i(x) p^{i-s-1}, \tag{2.5}$$

where $\xi_i(x) = \eta_{s+1-i}(x) \in \mathbb{F}_p$ for every $i = 1, \dots, s$. Using this representation, we can define an inner product on the space $\mathbb{Q}^n(p^s)$ as follows. For every $x, y \in \mathbb{Q}(p^s)$, we let

$$\langle x, y \rangle = \langle y, x \rangle = \sum_{i=1}^{s} \xi_i(x) \xi_{s+1-i}(y). \tag{2.6}$$

For vectors $X = (x_1, \dots, x_n)$ and $Y = (y_1, \dots, y_n)$ in $\mathbb{Q}^n(p^s)$, we write

$$\langle X, Y \rangle = \langle Y, X \rangle = \sum_{j=1}^{n} \langle x_j, y_j \rangle. \tag{2.7}$$

Remark. Instead of using the representation (2.5) and the definition (2.6), we could use the representation (2.3) and write $(x, y) = \eta_1(x)\eta_1(y) + \dots + \eta_s(x)\eta_s(y)$. In this case, we have $(x, y) = \langle x, \Phi y \rangle$, where Φ is a non-singular $s \times s$ matrix with entries in the finite field $\mathbb{F}_p$. Unfortunately, this matrix Φ would then be involved in all subsequent consideration.

Definition. Suppose that $D \subseteqq \mathbb{Q}^n(p^s)$ is a linear distribution, where $s \in \mathbb{N}_0$. The dual distribution $D^\perp \subseteqq \mathbb{Q}^n(p^s)$ is defined by

$$D^\perp = \{X \in \mathbb{Q}^n(p^s) \colon \langle X, Y \rangle = 0 \text{ for every } Y \in D\}.$$

It is easy to check that $D^\perp$ is a subspace of $\mathbb{Q}^n(p^s)$, and is therefore also a linear distribution. Furthermore, we have $(D^\perp)^\perp = D$, so that D and $D^\perp$ are mutually dual subspaces of $\mathbb{Q}^n(p^s)$, with the sum of the dimensions equal to ns. It follows immediately that $\#(D)\#(D^\perp) = p^{ns}$.

Lemma 2A. *Suppose that $D, D^\perp \subseteqq \mathbb{Q}^n(p^s)$ are mutually dual linear distributions. Then for every $A = (a_1, \dots, a_n) \in \mathbb{N}_0^n$ satisfying $0 \leqq a_j \leqq s$ for every $j = 1, \dots, n$, we have*

$$\#(D \cap \Delta_A^0) = p^{-a_1 - \cdots - a_n} \#(D) \#(D^\perp \cap \Delta_{A^*}^0),$$

where $A^* = (a_1^*, \dots, a_n^*) \in \mathbb{N}_0^n$ satisfies $a_j + a_j^* = s$ for every $j = 1, \dots, n$.

We shall establish this result, which is Lemma 4.3 in [26], in Section 7.

We next introduce two metrics on the vector space $\mathbb{Q}^n(p^s)$.

Recall that any $x \in \mathbb{Q}(p^s)$ can be written in the form (2.5), where $\xi_i(x) \in \mathbb{F}_p$ for every $i = 1, \dots, s$. By the Hamming weight $\varkappa(x)$, we mean the number of non-zero coefficients $\xi_i(x)$ in the representation (2.5); see [14]. We also define a non-Hamming weight by writing $\rho(0) = 0$ and writing

$$\rho(x) = \max\{i = 1, \dots, s\colon \xi_i(x) \neq 0\} \tag{2.8}$$

for every $x \neq 0$. For $X = (x_1, \dots, x_n) \in \mathbb{Q}^n(p^s)$, we now let

$$\varkappa(X) = \sum_{j=1}^n \varkappa(x_j) \quad \text{and} \quad \rho(X) = \sum_{j=1}^n \rho(x_j). \tag{2.9}$$

It is easy to check that $\varkappa(X) = \rho(X) = 0$ if and only if $X = 0$. One can also easily check the triangle inequalities for both weights. These give rise to metrics (or distances) on the vector space $\mathbb{Q}^n(p^s)$.

We remark here that the metric ρ was first introduced in the context of coding theory by Rosenbloom and Tsfasman [19], and in the context of uniform distribution independently by Martin and Stinson [15] and by Skriganov [26]. The interested reader is referred to [26] for further discussion.

Here we establish the following property of ρ which we shall need in Section 6.

Lemma 2B. *Suppose that $Y = (y_1, \dots, y_n) \in \mathbb{Q}^n(p^s)$ satisfies $\rho(y_j) \leqq a_j$, where $0 \leqq a_j \leqq s$ for every $j = 1, \dots, n$. Then $Y \in \Delta_{A^*}^0$, where $A^* = (a_1^*, \dots, a_n^*) \in \mathbb{N}_0^n$ satisfies $a_j + a_j^* = s$ for every $j = 1, \dots, n$.*

Proof. Simply note that for every $j = 1, \dots, n$, we have

$$y_j = \sum_{i=1}^{a_j} \xi_i(y_j) p^{i-s-1} \leqq (p-1) \sum_{i=1}^{a_j} p^{i-s-1} < p^{a_j - s} = p^{-a_j^*}. \quad \square$$

Suppose now that a linear distribution $D \subseteqq \mathbb{Q}^n(p^s)$ contains at least two points. We shall consider a Hamming weight

$$\varkappa(D) = \min\{\varkappa(X): X \in D\backslash\{0\}\}, \tag{2.10}$$

and a non-Hamming weight

$$\rho(D) = \min\{\rho(X): X \in D\backslash\{0\}\}. \tag{2.11}$$

Lemma 2C. *Suppose that $D, D^\perp \subset \mathbb{Q}^n(p^s)$ are mutually dual linear distributions of dimensions s and $(n-1)s$ respectively. Then the following statements are equivalent*:

(i) *D is an (n, s, δ)-net.*

(ii) *The non-Hamming weight $\rho(D^\perp) \geqq s + 1 - \delta$.*

This is a special case of Theorem 4.2 in [26]. It shows that the points of a linear distribution D are uniformly distributed in the cube U^n if the points of the dual distribution $D^\perp$ are well spaced with respect to the metric ρ. We shall give a proof of this in Section 7.

The main part of this paper is concerned with establishing the following result.

Lemma 2D. *Suppose that $D, D^\perp \subset \mathbb{Q}^n(p^s)$ are mutually dual linear distributions of dimensions s and $(n-1)s$ respectively. Suppose further that*

$$\varkappa(D^\perp) \geqq 2n + 1 \quad \textit{and} \quad \rho(D^\perp) \geqq s + 1 - \delta,$$

where δ is an integer satisfying $0 \leqq \delta \leqq s$. Then

$$\mathscr{L}_2[D] < 2^{n+1} p^{n+\delta} (s+1)^{\frac{1}{2}(n-1)}.$$

Remark. To establish our Theorem, it is sufficient to consider the special case when $\delta = 0$. However, we shall establish Lemma 2D for arbitrary δ.

The existence of distributions that satisfy the hypotheses of Lemma 2D above is established in the following way.

Let $\mathbb{F}_p[z]$ denote the ring of polynomials with coefficients in $\mathbb{F}_p$, the finite field of residues modulo p. Consider any polynomial

$$f(z) = \sum_{i=0}^{t-1} f_i z^i$$

in $\mathbb{F}_p[z]$, where $t = \deg f + 1$. For every integer $j \geqq 1$, we shall consider the j-th hyperderivative

$$\partial^j f(z) = \sum_{i=0}^{t-1} \binom{i}{j} f_i z^{i-j},$$

where $\binom{i}{j}$ denotes a binomial coefficient modulo p, with the usual convention that $\binom{i}{j} = 0$ when $j > i$. For more details, see Section 6.4 of [13].

Suppose that the prime p satisfies the condition $p \geqq gn$, where $g \in \mathbb{N}$ is fixed. Then there exist gn distinct residues $\beta_{i,\ell} \in \mathbb{F}_p$ with $1 \leqq i \leqq n$ and $1 \leqq \ell \leqq g$. For every $\sigma \in \mathbb{N}$ and $f \in \mathbb{F}_p[z]$, let

$$X(f) = \big(x_1(f), \dots, x_n(f)\big),$$

where for every $i = 1, \dots, n$,

$$x_i(f) = \sum_{\ell=1}^{g} p^{-(\ell-1)\sigma} \sum_{j=1}^{\sigma} \partial^{j-1} f(\beta_{i,\ell}) p^{-j}.$$

It is easy to see that the set

$$D(g,\sigma) = \{X(f) : f \in \mathbb{F}_p[z] \text{ and } \deg f < g\sigma\} \tag{2.12}$$

is a set of $p^{g\sigma}$ points in U^n.

Lemma 2E. *For every $g, \sigma \in \mathbb{N}$ satisfying $p \geqq gn$, the set $D(g,\sigma) \subset \mathbb{Q}^n(p^{g\sigma})$ is a linear distribution of dimension $g\sigma$. Furthermore, the dual linear distribution $\big(D(g,\sigma)\big)^{\perp} \subset \mathbb{Q}^n(p^{g\sigma})$ is of dimension $(n-1)g\sigma$, with*

$$\varkappa\big(\big(D(g,\sigma)\big)^{\perp}\big) \geqq g+1 \quad \textit{and} \quad \rho\big(\big(D(g,\sigma)\big)^{\perp}\big) \geqq g\sigma + 1.$$

This is a consequence of variants of Lemma 5.1 and Theorem 6.1 in [26]. We shall discuss this in Section 8. We remark also that the set (2.12) in the special case $g = 1$ was studied earlier by Faure [8].

3. Proof of the Theorem

Let $g = 2n$, and let $p \geqq gn = 2n^2$ be a prime. Given any natural number $N > 1$, we choose $\sigma \in \mathbb{N}$ such that

$$p^{g(\sigma-1)} < N \leqq p^{g\sigma}.$$

By Lemma 2E, the set $D(g,\sigma)$ of $p^{g\sigma}$ points in U^n is a linear distribution, and satisfies the hypotheses of Lemma 2D with $s = g\sigma$ and $\delta = 0$, and so

$$\mathscr{L}_2[D(g,\sigma)] < 2^{n+1} p^n (g\sigma+1)^{\frac{1}{2}(n-1)}.$$

It also follows from Lemma 2C that $D(g,\sigma)$ is an $(n, g\sigma, 0)$-net, so that the subset

$$\mathscr{D}_N^*(g) = D(g,\sigma) \cap \big([0, Np^{-g\sigma}) \times [0,1)^{n-1}\big)$$

contains exactly N points. Now let

$$\mathscr{D}_N = \mathscr{D}_N(g) = \{(N^{-1}p^{g\sigma}x_1, x_2, \dots, x_n) : (x_1, x_2, \dots, x_n) \in \mathscr{D}_N^*(g)\}.$$

Then it is easy to see that

$$\begin{aligned}\int_{U^n} |\mathscr{L}[\mathscr{D}_N(g); Y]|^2 \, \mathrm{d}Y &= N^{-1} p^{g\sigma} \int_{[0, Np^{-g\sigma}) \times [0,1)^{n-1}} |\mathscr{L}[D(g,\sigma); Y]|^2 \, \mathrm{d}Y \\ &\leqq N^{-1} p^{g\sigma} \int_{U^n} |\mathscr{L}[D(g,\sigma); Y]|^2 \, \mathrm{d}Y \\ &< N^{-1} p^{g\sigma} 2^{2n+2} p^{2n} (g\sigma + 1)^{n-1} \\ &< p^g 2^{2n+2} p^{2n} \left(\frac{\log N}{\log p} + g + 1 \right)^{n-1} \\ &< 2^{2n+2} p^{4n} (\log N + 2n + 1)^{n-1}.\end{aligned}$$

The Theorem now follows on taking square roots.

4. Walsh functions

First of all, we consider p-ary representation of integers $\ell \in \mathbb{N}_0$. Observe that any $\ell \in \mathbb{N}_0$ can be written uniquely in the form

$$\ell = \sum_{i=1}^{\infty} \lambda_i(\ell) p^{i-1}, \tag{4.1}$$

where the coefficients $\lambda_i(\ell) \in \mathbb{F}_p = \{0, 1, \ldots, p-1\}$ for every $i \in \mathbb{N}$. By the Hamming weight $\varkappa(\ell)$, we mean the number of non-zero coefficients $\lambda_i(\ell)$ in (4.1). We also define a non-Hamming weight by writing $\rho(0) = 0$ and writing

$$\rho(\ell) = \max\{i \in \mathbb{N} \colon \lambda_i(\ell) \neq 0\}$$

for every $\ell \in \mathbb{N}$. Note that for every $\ell \in \mathbb{N}$, we have

$$p^{\rho(\ell)-1} \leqq \ell < p^{\rho(\ell)}.$$

For $L = (\ell_1, \ldots, \ell_n) \in \mathbb{N}_0^n$, we now let

$$\varkappa(L) = \sum_{j=1}^{n} \varkappa(\ell_j) \quad \text{and} \quad \rho(L) = \sum_{j=1}^{n} \rho(\ell_j).$$

Consider the set

$$\mathbb{N}_0^n(p^s) = \{L = (\ell_1, \ldots, \ell_n) \in \mathbb{N}_0^n \colon 0 \leqq \ell_j < p^s \text{ for every } j = 1, \ldots, n\}.$$

For any two vectors $L = (\ell_1, \ldots, \ell_n)$ and $K = (k_1, \ldots, k_n)$ in $\mathbb{N}_0^n(p^s)$ and any two scalars $\alpha, \beta \in \mathbb{F}_p$, we write

$$\alpha L \oplus \beta K = (\alpha \ell_1 \oplus \beta k_1, \ldots, \alpha \ell_n \oplus \beta k_n) \in \mathbb{N}_0^n(p^s) \tag{4.2}$$

by setting

$$\lambda_i(\alpha\ell_j \oplus \beta k_j) = \alpha\lambda_i(\ell_j) + \beta\lambda_i(k_j) \pmod p$$

for every $i = 1, \dots, s$ and $j = 1, \dots, n$. It is easy to see that with respect to the arithmetic operations (4.2), each set $\mathbb{N}_0^n(p^s)$ forms a vector space of dimension ns over $\mathbb{F}_p$, while the set $\mathbb{N}_0^n$ forms an infinite dimensional vector space over $\mathbb{F}_p$.

We observe also that the mapping

$$\theta\colon \mathbb{Q}^n(p^s) \to \mathbb{N}_0^n(p^s)\colon (x_1, \dots, x_n) \mapsto (p^s x_1, \dots, p^s x_n) \tag{4.3}$$

is clearly an isomorphism of vector spaces. Furthermore, this isomorphism preserves the metrics $\varkappa$ and ρ. More precisely, for every $X \in \mathbb{Q}^n(p^s)$, we have

$$\varkappa(X) = \varkappa\big(\theta(X)\big) \quad \text{and} \quad \rho(X) = \rho\big(\theta(X)\big).$$

For every $\ell \in \mathbb{N}_0$ and every $x \in [0, 1)$, we let

$$w_\ell(x) = e_p\left(\sum_{i=1}^{\infty} \lambda_i(\ell)\eta_i(x)\right),$$

where $e_p(z) = \mathrm{e}^{2\pi \mathrm{i} z/p}$ for every real number z, and where the coefficients $\lambda_i(\ell)$ and $\eta_i(x)$ are given by (4.1) and (2.3) respectively. The functions w_ℓ are known as the Walsh functions if $p = 2$ and the Chrestenson or Chrestenson-Levy functions if $p > 2$. For simplicity, we refer to them all as Walsh functions here. A detailed study of such functions can be found in [11] or [23].

It is easy to see that $w_0(x) = 1$ for every $x \in \Delta_0^0 = [0, 1)$. Furthermore, for every $\ell \in \mathbb{N}$ satisfying $p^a \leqq \ell < p^{a+1}$ where $a \in \mathbb{N}_0$, $w_\ell(x)$ is constant and equal to a p-th root of unity on each interval Δ_{a+1}^m, where $0 \leqq m < p^{a+1}$. Furthermore, elementary intervals of the type Δ_{a+1}^m are maximal where $w_\ell(x)$ is constant.

The Walsh functions in higher dimensions are defined as follows.

For every $L = (\ell_1, \dots, \ell_n) \in \mathbb{N}_0^n$ and every $X = (x_1, \dots, x_n) \in U^n$, we let

$$W_L(X) = \prod_{j=1}^{n} w_{\ell_j}(x_j).$$

It is well known that

$$W_{L \oplus K}(X) = W_L(X) W_K(X)$$

for every $X \in U^n$ and $L, K \in \mathbb{N}_0^n$, and that

$$W_L(X \oplus Y) = W_L(X) W_L(Y)$$

for every $X, Y \in \mathbb{Q}^n(p^\infty)$ and $L \in \mathbb{N}_0^n$. Indeed, Walsh functions restricted to points $X \in \mathbb{Q}^n(p^s)$ are the additive characters of the vector space $\mathbb{Q}^n(p^s)$. For more details, see Chapter 5 of [13].

Furthermore, one can show that

$$W_{\theta(Y)}(X) = e_p(\langle Y, X\rangle)$$

for every $X, Y \in \mathbb{Q}^n(p^s)$, where the vector $\theta(Y)$ is defined by the mapping (4.3) and the inner product $\langle Y, X\rangle$ is given by (2.7). In other words, each additive character of the vector space $\mathbb{Q}^n(p^s)$ can be obtained as the restriction of a Walsh function $W_L(X)$ on $\mathbb{Q}^n(p^s)$ for some $L \in \mathbb{N}_0^n(p^s)$.

It is well known that the Walsh functions form an orthonormal basis in the Hilbert space $L_2(U^n)$ of square-integrable functions on the n-dimensional unit cube U^n. More precisely, we have

$$\int_{U^n} W_K(X)\overline{W_L(X)}\,\mathrm{d}X = \begin{cases} 1 & \text{if } K = L, \\ 0 & \text{if } K \neq L. \end{cases}$$

For each $f \in L_2(U^n)$, we have the Fourier-Walsh expansion

$$f(X) \simeq \sum_{L \in \mathbb{N}_0^n} \tilde{f}_L \overline{W_L(X)}, \tag{4.4}$$

where the Fourier-Walsh coefficient

$$\tilde{f}_L = \int_{U^n} W_L(X) f(X)\,\mathrm{d}X.$$

Here the symbol $\simeq$ denotes that the series (4.4) converges in the L_2-norm. We also have the identity

$$\int_{U^n} |f(X)|^2\,\mathrm{d}X = \sum_{L \in \mathbb{N}_0^n} |\tilde{f}_L|^2.$$

Known results on characters of abelian groups can be restated in terms of Walsh functions. See [13] or [14].

Lemma 4A. *Suppose that $D, D^\perp \subseteqq \mathbb{Q}^n(p^s)$ are mutually dual linear distributions. Then for every $L \in \mathbb{N}_0^n(p^s)$, we have*

$$\sum_{X \in D} W_L(X) = \begin{cases} \#(D) & \text{if } L \in \theta(D^\perp), \\ 0 & \text{if } L \notin \theta(D^\perp), \end{cases}$$

where $\theta(D^\perp) = \{\theta(Y)\colon Y \in D^\perp\}$ denotes the image of $D^\perp$ under the mapping θ.

To handle Fourier-Walsh expansions pointwise, we consider the following. For more details, see Section 2.8 of [11].

Suppose that $f \in L_2\big([0,1)\big)$ is given. For every $s \in \mathbb{N}_0$, for every $m \in \mathbb{N}_0$ satisfying $0 \leqq m < p^s$, and for every $x \in \Delta_s^m = \big[mp^{-s}, (m+1)p^{-s}\big)$, we let

$$f_s(x) = p^s \int_{\Delta_s^m} f(t)\,\mathrm{d}t. \tag{4.5}$$

It is clear that if f is real valued, then for every $x \in \Delta_s^m$, we have

$$\inf\{f(x)\colon x \in \Delta_s^m\} \leqq f_s(x) \leqq \sup\{f(x)\colon x \in \Delta_s^m\}. \tag{4.6}$$

Lemma 4B. *Suppose that $f \in L_2\big([0,1)\big)$. Then for every $s \in \mathbb{N}_0$, the function f_s has the finite Fourier-Walsh expansion*

$$f_s(x) = \sum_{\ell=0}^{p^s-1} \tilde{f}_\ell \overline{w_\ell(x)},$$

where the coefficients

$$\tilde{f}_\ell = \int_0^1 w_\ell(x) f(x)\,\mathrm{d}x$$

are the Fourier-Walsh coefficients of f.

5. Preparation for the proof of Lemma 2D

For any $Y = (y_1, \dots, y_n) \in U^n$, we consider the characteristic function $\chi(Y, X)$ of the rectangular box $B_Y = [0, y_1) \times \cdots \times [0, y_n)$, so that

$$\chi(Y, X) = \begin{cases} 1 & \text{if } X \in B_Y, \\ 0 & \text{if } X \notin B_Y. \end{cases}$$

It is clear that if $X = (x_1, \dots, x_n)$, then

$$\chi(Y, X) = \prod_{j=1}^{n} \chi(y_j, x_j),$$

where for every $j = 1, \dots, n$,

$$\chi(y_j, x_j) = \begin{cases} 1 & \text{if } x_j \in [0, y_j), \\ 0 & \text{if } x_j \notin [0, y_j). \end{cases} \tag{5.1}$$

Our task is to find the Fourier-Walsh expansions of the characteristic functions.

For $y \in [0,1)$, let $\Delta_s(y) = \big[mp^{-s}, (m+1)p^{-s}\big)$ be the unique elementary interval of length p^{-s} containing y, and let $\varepsilon_s(y, x)$ denote the characteristic function of this interval $\Delta_s(y)$, so that

$$\varepsilon_s(y, x) = \begin{cases} 1 & \text{if } x \in \Delta_s(y), \\ 0 & \text{if } x \notin \Delta_s(y). \end{cases}$$

Lemma 5A. *For every $s \in \mathbb{N}$, we have*

$$\chi(y, x) = \chi_s(y, x) + r_s(y, x), \tag{5.2}$$

where $\chi_s(y,x)$ is a piecewise constant function of $x \in [0,1)$ with the finite Fourier-Walsh expansion

$$\chi_s(y,x) = \sum_{\ell=0}^{p^s-1} \tilde{\chi}_\ell(y)\overline{w_\ell(x)} = y + \sum_{\ell=1}^{p^s-1} \tilde{\chi}_\ell(y)\overline{w_\ell(x)}, \tag{5.3}$$

with $\tilde{\chi}_0(y) = y$ and

$$\tilde{\chi}_\ell(y) = \int_0^y w_\ell(x)\,\mathrm{d}x \tag{5.4}$$

whenever $\ell \geqq 1$, and where

$$0 \leqq \chi_s(y,x) \leqq 1 \quad \textit{and} \quad 0 \leqq r_s(y,x) \leqq \varepsilon_s(y,x). \tag{5.5}$$

In particular, $r_s(y,x) = 0$ whenever $x \notin \Delta_s(y)$.

Proof. Let $f(x) = \chi(y,x)$, and let $\chi_s(y,x) = f_s(x)$, where $f_s(x)$ is given by (4.5). The relations (5.3) and (5.4) follow immediately from Lemma 4B, while the inequalities (5.5) follow immediately from the inequalities (4.6). □

Suppose that $Y = (y_1, \ldots, y_n)$ and $X = (x_1, \ldots, x_n)$ are vectors in U^n. For every $j = 1, \ldots, n$ and every $s \in \mathbb{N}$, let $\varepsilon_{j,s}(Y,X) = \varepsilon_s(y_j, x_j)$ denote the characteristic function of the elementary box

$$\Delta_{j,s}(Y) = [0,1)^{j-1} \times \Delta_s(y_j) \times [0,1)^{n-j}.$$

Furthermore, let

$$\chi_s(Y,X) = \prod_{j=1}^{n} \chi_s(y_j, x_j), \tag{5.6}$$

where the terms $\chi_s(y_j, x_j)$ are given by (5.3) and (5.4).

Lemma 5B. *For every $s \in \mathbb{N}$, we have*

$$\chi(Y,X) = \chi_s(Y,X) + \sum_{j=1}^{n} r_{j,s}(Y,X), \tag{5.7}$$

where

$$0 \leqq r_{j,s}(Y,X) \leqq \varepsilon_{j,s}(Y,X) \tag{5.8}$$

for every $j = 1, \ldots, n$. In particular, $r_{j,s}(Y,X) = 0$ whenever $X \notin \Delta_{j,s}(Y)$.

Proof. We proceed by induction on n. The case $n = 1$ is given by Lemma 5A. Assume next that the assertion holds for a particular value of n. Write $X = (x_1, \ldots, x_n)$ and

$Y = (y_1, \dots, y_n)$, and let $X^+ = (X, x_{n+1})$ and $Y^+ = (Y, y_{n+1})$ be in U^{n+1}. Then it follows from (5.6), (5.7) and (5.2) that

$$\begin{aligned}
\chi(Y^+, X^+) &= \chi(Y, X)\chi(y_{n+1}, x_{n+1}) \\
&= \left(\chi_s(Y, X) + \sum_{j=1}^{n} r_{j,s}(Y, X)\right)\chi(y_{n+1}, x_{n+1}) \\
&= \chi_s(Y, X)\chi(y_{n+1}, x_{n+1}) + \sum_{j=1}^{n} r_{j,s}(Y, X)\chi(y_{n+1}, x_{n+1}) \\
&= \chi_s(Y, X)\chi_s(y_{n+1}, x_{n+1}) + \chi_s(Y, X)r_s(y_{n+1}, x_{n+1}) \\
&\quad + \sum_{j=1}^{n} r_{j,s}(Y, X)\chi(y_{n+1}, x_{n+1}) \\
&= \chi_s(Y^+, X^+) + \sum_{j=1}^{n+1} r_{j,s}(Y^+, X^+),
\end{aligned}$$

where

$$r_{n+1,s}(Y^+, X^+) = \chi_s(Y, X)r_s(y_{n+1}, x_{n+1})$$

and

$$r_{j,s}(Y^+, X^+) = r_{j,s}(Y, X)\chi(y_{n+1}, x_{n+1})$$

for every $j = 1, \dots, n$. Using (5.1), (5, 5), (5.6) and (5.8), we conclude that for every $j = 1, \dots, n+1$, we have

$$0 \leqq r_{j,s}(Y^+, X^+) \leqq \varepsilon_{j,s}(Y^+, X^+). \quad \square$$

We now return to Lemma 5A. The evaluation of the integral (5.4) is a very interesting problem, and was studied by Fine [9] for $p = 2$ and by Price [18] for $p > 2$.

For every $\ell \in \mathbb{N}$, let $\tau(\ell) \in \mathbb{N}_0$ be given by the truncation

$$\tau(\ell) = \sum_{i=1}^{\rho(\ell)-1} \lambda_i(\ell) p^{i-1}, \tag{5.9}$$

where the largest term in the p-ary representation of ℓ has been removed. Note that

$$\ell = \lambda_{\rho(\ell)}(\ell) p^{\rho(\ell)-1} + \tau(\ell).$$

We have the Fine-Price formula, that for every $\ell \in \mathbb{N}$,

$$p^{\rho(\ell)} \int_0^y w_\ell(x)\,\mathrm{d}x = (1 - \zeta^{\lambda(\ell)})^{-1} w_{\tau(\ell)}(y) + \varphi_\ell(y), \tag{5.10}$$

where $\lambda(\ell) = \lambda_{p(\ell)}(\ell)$, where $\zeta = e_p(1)$ is a primitive p-th root of unity, and where

$$\varphi_\ell(y) = \left(\frac{1}{2} - (1-\zeta^{\lambda(\ell)})^{-1}\right) w_\ell(y) + \sum_{i=1}^{\infty} p^{-i} \sum_{j=1}^{p-1} \zeta^j (1-\zeta^j)^{-1} w_{\ell + jp^{\rho(\ell)+i-1}}(y). \tag{5.11}$$

Note that for $j = 1, \dots, p-1$, we have

$$|1-\zeta^j| = 2\left|\sin\frac{\pi j}{p}\right| \geqq 2\sin\frac{\pi}{p} \geqq \frac{4}{p},$$

so that the series in (5.11) converges absolutely for every $y \in [0,1)$. Note also that

$$|1-\zeta^{\lambda(\ell)}| \geqq \frac{4}{p}$$

by a similar argument, since $\lambda(\ell) \not\equiv 0 \pmod p$.

It is convenient to write $\tau(0) = 0$, so that the expression (5.9) remains valid.

Definition. Suppose that $\mathscr{H}$ is a complex Hilbert space with norm $\|\cdot\|_{\mathscr{H}}$. Suppose also that A is a finite or countable index set. We say that a subset $\{\psi_\alpha : \alpha \in A\} \subset \mathscr{H}$ is quasi-orthonormal if

$$\left\| \sum_{\alpha \in A} c_\alpha \psi_\alpha \right\|_{\mathscr{H}}^2 \leqq \sum_{\alpha \in A} |c_\alpha|^2$$

for all square-summable complex sequences $\{c_\alpha : \alpha \in A\}$.

Lemma 5C. *Suppose that $\psi_0(y) = y$ and $\psi_\ell(y) = p^{-1}\varphi_\ell(y)$ for every $\ell \in \mathbb{N}$. Then the set $\{\psi_\ell : \ell \in \mathbb{N}_0\} \subset L_2\big([0,1)\big)$ is quasi-orthonormal.*

Proof. For every $\ell \in \mathbb{N}$, we can write

$$\varphi_\ell(y) = \sum_{k=1}^{\infty} Q_{k,\ell} w_k(y),$$

where

$$Q_{k,\ell} = \left(\frac{1}{2} - (1-\zeta^{\lambda(\ell)})^{-1}\right)\delta(k,\ell) + \sum_{i=1}^{\infty} p^{-i} \sum_{j=1}^{p-1} \zeta^j (1-\zeta^j)^{-1} \delta(k, \ell + jp^{\rho(\ell)+i-1}).$$

Here the Kronecker function $\delta(k,\ell) = 1$ if $k = \ell$ and $\delta(k,\ell) = 0$ if $k \neq \ell$. Orthonormality of the Walsh functions in $L_2\big([0,1)\big)$ implies that

$$\left\| \sum_{\ell=1}^{\infty} c_\ell \varphi_\ell \right\|_{L_2}^2 \leqq Q_1 Q_2 \sum_{\ell=1}^{\infty} |c_\ell|^2$$

for all square-summable complex sequences $\{c_\ell : \ell \in \mathbb{N}\}$, where

$$Q_1 = \sup_{\ell \in \mathbb{N}} \sum_{k=1}^{\infty} |Q_{k,\ell}| \quad \text{and} \quad Q_2 = \sup_{k \in \mathbb{N}} \sum_{\ell=1}^{\infty} |Q_{k,\ell}|.$$

For details, see Theorem 6 of Section 2.5 of [2]. For any fixed $i, k \in \mathbb{N}$ and any $j = 1, \dots, p-1$, there is at most one value $\ell \in \mathbb{N}$ such that $\ell + jp^{\rho(\ell)+i-1} = k$, in view of the uniqueness of p-ary representations. Hence

$$Q_1, Q_2 \leqq \frac{1}{2} + \frac{p}{4} + \sum_{i=1}^{\infty} p^{-i} \sum_{j=1}^{p-1} |1 - \zeta^j|^{-1} \leqq \frac{1}{2} + \frac{p}{4} + \frac{(p-1)p}{4} \sum_{i=1}^{\infty} p^{-i} = \frac{p+1}{2}.$$

It follows that as long as $p \geqq 3$, we have

$$\left\| \sum_{\ell=0}^{\infty} c_\ell \psi_\ell \right\|_{L_2}^2 \leqq 2 \left(|c_0|^2 \|\psi_0\|_{L_2}^2 + \left\| \sum_{\ell=1}^{\infty} c_\ell \psi_\ell \right\|_{L_2}^2 \right) = 2 \left(\frac{1}{3} |c_0|^2 + \frac{1}{p^2} \left\| \sum_{\ell=1}^{\infty} c_\ell \varphi_\ell \right\|_{L_2}^2 \right)$$

$$\leqq \frac{2}{3} |c_0|^2 + \frac{(p+1)^2}{2p^2} \sum_{\ell=1}^{\infty} |c_\ell|^2 \leqq \sum_{\ell=0}^{\infty} |c_\ell|^2. \quad \square$$

We remark that we are not trying to find the best numerical constants in our estimates.

We next return to Lemma 5B.

For every $L = (\ell_1, \dots, \ell_n) \in \mathbb{N}_0^n$, $Y = (y_1, \dots, y_n) \in U^n$ and $J \subseteqq \{1, \dots, n\}$, let

$$\Psi_L^{(J)}(Y) = \left(\prod_{j \notin J} \psi_{\ell_j}(y_j) \right) \left(\prod_{j \in J} w_{\ell_j}(y_j) \right). \tag{5.12}$$

Lemma 5D. *For every $J \subseteqq \{1, \dots, n\}$, the set*

$$\{\Psi_L^{(J)} : L \in \mathbb{N}_0^n\} \subset L_2(U^n)$$

is quasi-orthonormal.

Proof. The Hilbert space $L_2(U^n)$ is a tensor product of n copies of the Hilbert space $L_2\big([0,1)\big)$. The functions (5.12) are also tensor products of the functions ψ_ℓ and w_ℓ. The result follows from Lemma 5C and the orthonormality of the Walsh functions. $\square$

To use Lemma 5D, we need to consider the truncation of integer vectors. For every $L = (\ell_1, \dots, \ell_n) \in \mathbb{N}_0^n$ and $J \subseteqq \{1, \dots, n\}$, let $\tau_J(L) = (k_1, \dots, k_n)$, where

$$k_j = \begin{cases} \tau(\ell_j) & \text{if } j \in J, \\ \ell_j & \text{if } j \notin J. \end{cases}$$

Lemma 5E. *For every $s \in \mathbb{N}$, we have the Fourier-Walsh expansion*

$$\chi_s(Y,X) = y_1 \dots y_n + \sum_{L \in \mathbb{N}_0^n(p^s)\setminus\{0\}} p^{-\rho(L)} V_L(Y)\,\overline{W_L(X)}, \tag{5.13}$$

where

$$V_L(Y) = \sum_J c_{J,L}\Psi^{(J)}_{\tau_J(L)}(Y), \tag{5.14}$$

with complex coefficients $c_{J,L}$ satisfying the bound

$$|c_{J,L}| \leqq p^n, \tag{5.15}$$

and where the summation in (5.14) *is extended over all subsets $J \subseteqq \{1,\dots,n\}$.*

Proof. Note from (5.4) and (5.10) that for every $\ell \in \mathbb{N}$, we can write

$$\tilde{\chi}_\ell(y) = p^{-\rho(\ell)}\big(a_\ell w_{\tau(\ell)}(y) + b_\ell \psi_\ell(y)\big), \tag{5.16}$$

where $a_\ell = (1-\zeta^{\lambda(\ell)})^{-1}$ and $b_\ell = p$ for every $\ell \in \mathbb{N}$. The identity (5.16) can be extended to include the case $\ell = 0$ if we take $a_0 = 0$ and $b_0 = 1$. Clearly

$$|a_\ell| \leqq p \quad \text{and} \quad |b_\ell| \leqq p \tag{5.17}$$

for every $\ell \in \mathbb{N}_0$. On the other hand, it follows from (5.6), (5.3) and (5.16) that

$$\chi_s(Y,X) = \prod_{j=1}^n \left(\sum_{\ell_j=0}^{p^s-1} p^{-\rho(\ell_j)}\big(a_{\ell_j} w_{\tau(\ell_j)}(y_j) + b_{\ell_j}\psi_{\ell_j}(y_j)\big)\overline{w_{\ell_j}(x_j)} \right).$$

We now multiply out, and note that each coefficient $c_{J,L}$ is a product of n numbers, each of which is equal to a_{ℓ_j} or b_{ℓ_j}. The desired result now follows on noting the inequalities (5.17). □

We complete this section with a simple but crucial result on the relationship between the truncated integer vectors $\tau_J(L)$ and the values $\varkappa(L)$ given by the Hamming metric. The reader familiar with coding theory will see easily that the arguments in the proofs of Lemma 5F and of Lemma 6B below are typical of methods related to error-correcting codes. The interested reader is referred to [14].

Lemma 5F. (i) *Suppose that $\tau_J(L) = 0$, where $L \in \mathbb{N}_0^n$. Then $\varkappa(L) \leqq \#(J) \leqq n$.*

(ii) *Suppose that $\tau_J(L) = \tau_J(L')$, where $L, L' \in \mathbb{N}_0^n$. Then $\varkappa(L \ominus L') \leqq 2\#(J) \leqq 2n$, where $\ominus$ denotes subtraction related to the operation $\oplus$ in $\mathbb{N}_0^n$.*

Proof. It is easy to see from (5.9) and $\tau(0) = 0$ that the inequality $\varkappa(\ell) \leqq 1$ holds for every $\ell \in \mathbb{N}_0$ satisfying $\tau(\ell) = 0$, and that the inequality $\varkappa(\ell \ominus \ell') \leqq 2$ holds for every $\ell, \ell' \in \mathbb{N}_0$ satisfying $\tau(\ell) = \tau(\ell')$. □

6. Proof of Lemma 2D

Suppose that $D, D^\perp \subseteqq \mathbb{Q}^n(p^s)$ are mutually dual linear distributions of dimensions s and $(n-1)s$ respectively. Suppose further that $\varkappa(D^\perp) \geqq 2n+1$ and $\rho(D^\perp) \geqq s+1-\delta$. Recall that $\#(D) = p^s$ and $\#(D^\perp) = p^{(n-1)s}$.

Using Lemma 5B, we can write

$$\mathscr{L}[D; Y] = \mathscr{M}[D; Y] + \sum_{j=1}^{n} R_j[D; Y],$$

where

$$\mathscr{M}[D; Y] = \sum_{X \in D} \chi_s(Y, X) - p^s y_1 \dots y_n = \sum_{X \in D} \big(\chi_s(Y, X) - y_1 \dots y_n\big), \tag{6.1}$$

and where, for every $j = 1, \dots, n$,

$$R_j[D; Y] = \sum_{X \in D} r_{j,s}(Y, X).$$

Lemma 6A. *Suppose that $\rho(D^\perp) \geqq s+1-\delta$. Then for every $j = 1, \dots, n$ and every $Y \in U^n$, we have $0 \leqq R_j[D; Y] \leqq p^\delta$.*

Proof. To see this, note that it follows from Lemma 5B that

$$0 \leqq R_j[D; Y] \leqq \sum_{X \in D} \varepsilon_{j,s}(Y, X) = \#\big(D \cap \Delta_{j,s}(Y)\big).$$

The elementary box $\Delta_{j,s}(Y)$ has volume p^{-s} and is contained in some elementary box Δ_A^M of volume $p^{\delta-s}$. On the other hand, it follows from Lemma 2C that D is an (n, s, δ)-net, so that Δ_A^M contains precisely p^δ points of D. Hence

$$\#\big(D \cap \Delta_{j,s}(Y)\big) \leqq \#(D \cap \Delta_A^M) = p^\delta. \quad \square$$

It now follows that for every $Y \in U^n$, we have

$$|\mathscr{L}[D; Y]| \leqq |\mathscr{M}[D; Y]| + np^\delta.$$

Write

$$\mathscr{M}_2[D] = \left(\int_{U^n} |\mathscr{M}[D; Y]|^2 \, \mathrm{d}Y \right)^{1/2}.$$

Then

$$(\mathscr{L}_2[D])^2 \leqq 2(\mathscr{M}_2[D])^2 + 2n^2 p^{2\delta}. \tag{6.2}$$

Next, it follows from (6.1), (5.13), Lemma 4A and (5.14) that

$$\begin{aligned}\mathscr{M}[D;Y] &= \sum_{X\in D}\sum_{L\in\mathbb{N}_0^n(p^s)\setminus\{0\}} p^{-\rho(L)}V_L(Y)\overline{W_L(X)}\\ &= \sum_{L\in\mathbb{N}_0^n(p^s)\setminus\{0\}} p^{-\rho(L)}V_L(Y)\overline{\sum_{X\in D} W_L(X)}\\ &= \sum_{L\in\theta(D^\perp)\setminus\{0\}} p^{s-\rho(L)}V_L(Y)\\ &= \sum_J \sum_{L\in\theta(D^\perp)\setminus\{0\}} c_{J,L}p^{s-\rho(L)}\Psi^{(J)}_{\tau_J(L)}(Y)\\ &= \sum_J \sum_{K\in\mathbb{N}_0^n(p^s)} v_{J,K}\Psi_K^{(J)}(Y),\end{aligned}$$

where for every $K\in\mathbb{N}_0^n(p^s)$,

$$v_{J,K} = \sum_{L\in\Omega_{J,K}} c_{J,L}p^{s-\rho(L)}$$

with

$$\Omega_{J,K} = \{L\in\theta(D^\perp)\setminus\{0\}\colon \tau_J(L) = K\}.$$

If we write

$$u_{J,K} = \sum_{L\in\Omega_{J,K}} p^{s-\rho(L)},$$

then it follows from (5.15) that

$$|v_{J,K}| \leqq p^n u_{J,K}.$$

Note that there are exactly 2^n subsets $J\subseteqq\{1,\dots,n\}$. It follows from Lemma 5D that

$$\begin{aligned}(\mathscr{M}_2[D])^2 &\leqq 2^n\sum_J \int\limits_{U^n}\left|\sum_{K\in\mathbb{N}_0^n(p^s)} v_{J,K}\Psi_K^{(J)}(Y)\right|^2 \mathrm{d}Y\\ &\leqq 2^n\sum_J\sum_{K\in\mathbb{N}_0^n(p^s)}|v_{J,K}|^2 \leqq 2^n p^{2n}\sum_J U_J,\end{aligned}\tag{6.3}$$

where

$$U_J = \sum_{K\in\mathbb{N}_0^n(p^s)} u_{J,K}^2.$$

Lemma 6B. *Suppose that* $\varkappa(D^\perp)\geqq 2n+1$. *Then the following hold*:

(i) *Each subset* $\Omega_{J,0}$ *is empty*.

(ii) *Each subset* $\Omega_{J,K}$, *where* $K\in\mathbb{N}_0^n(p^s)$, *either is empty or consists of a single point* $L = L_{J,K}\in\theta(D^\perp)\setminus\{0\}$.

Proof. Suppose that $\Omega_{J,0}$ contains a point $L \in \theta(D^\perp)\backslash\{0\}$. Then $\varkappa(L) \leqq n$ by Lemma 5F(i). Write $L = \theta(Y)$, where $Y \in D^\perp\backslash\{0\}$. Then $\varkappa(Y) \leqq n$, contradicting the assumption that $\varkappa(D^\perp) \geqq 2n+1$. Suppose next that $\Omega_{J,K}$ contains two distinct points $L, L' \in \theta(D^\perp)\backslash\{0\}$. Then $\varkappa(L \ominus L') \leqq 2n$ by Lemma 5F(ii). Write $L = \theta(Y)$ and $L' = \theta(Y')$, where $Y, Y' \in D^\perp\backslash\{0\}$. Then $\varkappa(Y \ominus Y') \leqq 2n$. This means that there exists $Y'' \in D^\perp\backslash\{0\}$ such that $\varkappa(Y'') \leqq 2n$, again contradicting the assumption that $\varkappa(D^\perp) \geqq 2n+1$. □

Write

$$Q[D^\perp] = \sum_{Y \in D^\perp\backslash\{0\}} p^{2(s-\rho(Y))}.$$

Since the mapping θ is metric preserving, it follows that

$$Q[D^\perp] = \sum_{L \in \theta(D^\perp)\backslash\{0\}} p^{2(s-\rho(L))}.$$

Lemma 6C. *Suppose that $\varkappa(D^\perp) \geqq 2n+1$. Then for every subset $J \subseteqq \{1,\dots,n\}$, we have $U_J = Q[D^\perp]$. In particular, the value of U_J is independent of the choice of the subset $J \subseteqq \{1,\dots,n\}$.*

Proof. Write

$$\mathscr{E}_J = \{K \in \mathbb{N}_0^n(p^s)\colon \Omega_{J,K} \text{ is non-empty}\}.$$

In view of Lemma 6B, we have the following:

- $0 \notin \mathscr{E}_J$.

- For every $K \in \mathscr{E}_J$, there exists a unique element $L_{J,K} \in \theta(D^\perp)\backslash\{0\}$ such that $\Omega_{J,K} = \{L_{J,K}\}$.

- For every $L \in \theta(D^\perp)\backslash\{0\}$, there exists a unique element $K = K_{J,L} \in \mathscr{E}_J$ such that $\Omega_{J,K_{J,L}} = \{L\}$.

It follows that for every J, there is a bijection between the sets $\mathscr{E}_J$ and $\theta(D^\perp)\backslash\{0\}$. It is easy to see that

$$u_{J,K} = \begin{cases} p^{s-\rho(L_{J,K})} & \text{if } K \in \mathscr{E}_J, \\ 0 & \text{if } K \notin \mathscr{E}_J. \end{cases}$$

Hence

$$U_J = \sum_{K \in \mathscr{E}_J} p^{2(s-\rho(L_{J,K}))} = \sum_{L \in \theta(D^\perp)\backslash\{0\}} p^{2(s-\rho(L))} = Q[D^\perp]. \quad \square$$

Lemma 6D. *Suppose that $\rho(D^\perp) \geqq s+1-\delta$. Then $Q[D^\perp] < p^{2\delta}(s+1)^{n-1}$.*

Proof. We can write

$$Q[D^\perp] = \sum_{a_1+\cdots+a_n \geqq s+1-\delta} p^{2(s-a_1-\cdots-a_n)}\mu_A,$$

where for every $A = (a_1, \dots, a_n) \in \mathbb{N}_0^n$ with $0 \leqq a_j \leqq s$ for every $j = 1, \dots, n$,

$$\mu_A = \#\big(\{ Y = (y_1, \dots, y_n) \in D^\perp \colon \rho(y_j) = a_j \text{ for every } j = 1, \dots, n\}\big).$$

Using Lemma 2B and Lemma 2A, we have

$$\begin{aligned} \mu_A &\leqq \#\big(\{ Y = (y_1, \dots, y_n) \in D^\perp \colon \rho(y_j) \leqq a_j \text{ for every } j = 1, \dots, n\}\big) \\ &\leqq \#(D^\perp \cap \Delta^0_{A^*}) = p^{a_1 + \cdots + a_n - s} \#(D \cap \Delta^0_A). \end{aligned}$$

Note that the elementary box Δ^0_A has volume $p^{-a_1 - \cdots - a_n} \leqq p^{\delta - s - 1} < p^{\delta - s}$, so that it is contained in an elementary box of volume $p^{\delta - s}$. On the other hand, it follows from Lemma 2C that D is an (n, s, δ)-net. Hence $\#(D \cap \Delta^0_A) \leqq p^\delta$, and so we must have $\mu_A \leqq p^{a_1 + \cdots + a_n - s + \delta}$. It follows that

$$Q[D^\perp] \leqq p^\delta \sum_{a_1 + \cdots + a_n \geqq s + 1 - \delta} p^{s - a_1 - \cdots - a_n} = p^\delta \sum_{t = s + 1 - \delta}^{ns} p^{s - t} v_t,$$

where

$$v_t = \#\big(\{(a_1, \dots, a_n) \in \mathbb{N}_0^n \colon a_1 + \cdots + a_n = t \text{ and } a_j \leqq s \text{ for every } j = 1, \dots, n\}\big).$$

It is easy to show that $v_t \leqq (s+1)^{n-1}$ always. Hence

$$Q[D^\perp] \leqq p^\delta (s+1)^{n-1} \sum_{t = s + 1 - \delta}^{ns} p^{s - t} < p^{2\delta - 1} (s+1)^{n-1} \sum_{t=0}^{\infty} p^{-t} < p^{2\delta} (s+1)^{n-1}. \quad \square$$

Recall that there are precisely 2^n subsets $J \subseteqq \{1, \dots, n\}$. Combining (6.3), Lemma 6C and Lemma 6D, we have

$$(\mathscr{M}_2[D])^2 < 2^{2n} p^{2n + 2\delta} (s+1)^{n-1}.$$

It follows from (6.2) that

$$(\mathscr{L}_2[D])^2 < 2^{2n+1} p^{2n + 2\delta} (s+1)^{n-1} + 2n^2 p^{2\delta} < 2^{2n+2} p^{2n + 2\delta} (s+1)^{n-1},$$

and so

$$\mathscr{L}_2[D] < 2^{n+1} p^{n + \delta} (s+1)^{\frac{1}{2}(n-1)}.$$

This completes the proof of Lemma 2D.

7. Linear distributions

In this section and the next, we shall discuss ideas in [26], adapted to the context of our investigation here.

We begin by defining a special class of elementary boxes. For every $s \in \mathbb{N}_0$, we denote by $\mathfrak{E}_s$ the class of all elementary boxes of the type Δ^M_A defined by (2.1) but with the

extra restriction that the vector $A = (a_1, \dots, a_n)$ satisfies the condition $0 \leqq a_j \leqq s$ for every $j = 1, \dots, n$. In other words, $\mathfrak{E}_s$ denotes the class of all elementary boxes with side lengths at least p^{-s}.

Definition. Suppose that $s, k \in \mathbb{N}_0$ satisfy $0 \leqq k \leqq n$. A set $D \subset U^n$ of p^{ks} points is called an optimum $[n, k, s]$-distribution (in base p) if every elementary box in $\mathfrak{E}_s$ of volume p^{-ks} contains exactly one point of D.

Remark. It is easy to show that any optimum $[n, k, s]$-distribution (in base p) is also an $\big(n, ks, (k-1)s\big)$-net (in base p). The case $k = 1$ is of special interest. Any optimum $[n, 1, s]$-distribution (in base p) is also an $(n, s, 0)$-net (in base p).

Suppose that a subset $D \subseteqq \mathbb{Q}^n(p^s)$ contains at least two points. We shall define a Hamming weight $\varkappa(D)$ and a non-Hamming weight $\rho(D)$ as follows.

For any $X \in \mathbb{Q}^n(p^s)$, we define $\varkappa(X)$ and $\rho(X)$ as before by (2.9), and write

$$\varkappa(D) = \min\{\varkappa(X \ominus X') \colon X, X' \in D \text{ and } X \neq X'\} \tag{7.1}$$

and

$$\rho(D) = \min\{\rho(X \ominus X') \colon X, X' \in D \text{ and } X \neq X'\}. \tag{7.2}$$

It is not difficult to show that if $D \subseteqq \mathbb{Q}^n(p^s)$ is a linear distribution, then (7.1) and (7.2) are equivalent to (2.10) and (2.11) respectively.

Lemma 7A. *Suppose that $D \subseteqq \mathbb{Q}^n(p^s)$ is a set of p^{ks} points, where $0 \leqq k \leqq n$. Then the non-Hamming weight $\rho(D) \leqq (n-k)s + 1$.*

Proof. For any $X = (x_1, \dots, x_n) \in \mathbb{Q}^n(p^s)$, the projection

$$P_k X = (x_1, \dots, x_k, 0, \dots, 0) \in \mathbb{Q}^n(p^s)$$

clearly satisfies $X = P_k X \oplus (X \ominus P_k X)$, and

$$\rho(X) = \rho(P_k X) + \rho(X \ominus P_k X) \leqq \rho(P_k X) + (n-k)s$$

trivially. It follows that

$$\rho(D) \leqq (n-k)s + \min\{\rho(P_k X \ominus P_k X') \colon X, X' \in D \text{ and } X \neq X'\}.$$

Note that $\#\big(\{P_k X \colon X \in \mathbb{Q}^n(p^s)\}\big) = p^{ks} = \#(D)$. There are two possibilities. If $P_k X = P_k X'$ for some distinct $X, X' \in D$, then clearly $\rho(D) \leqq (n-k)s$. Alternatively, we must have $\{P_k X \colon X \in D\} = \{P_k X \colon X \in \mathbb{Q}^n(p^s)\}$, and clearly $\rho(D) \leqq (n-k)s + 1$ in this case. □

Suppose that $\Delta_A^M \in \mathfrak{E}_s$. Then it is easy to see that the intersection

$$V_A^M = \mathbb{Q}^n(p^s) \cap \Delta_A^M$$

is an affine subspace of the vector space $\mathbb{Q}^n(p^s)$, and that

$$V_A^M = V_A^0 \oplus Y_A^M, \tag{7.3}$$

where the vector $Y_A^M \in \mathbb{Q}^n(p^s)$ is defined uniquely up to translations in V_A^0.

Lemma 7B. (i) *For any $X \in \mathbb{Q}^n(p^s)$, we have $\rho(X) \leqq (n-k)s$ if and only if $X \in \Delta_A^0$ for some elementary box $\Delta_A^0 \in \mathfrak{E}_s$ with volume at most p^{-ks}.*

(ii) *For any $X, X' \in \mathbb{Q}^n(p^s)$, we have $\rho(X \ominus X') \leqq (n-k)s$ if and only if $X, X' \in \Delta_A^M$ for some elementary box $\Delta_A^M \in \mathfrak{E}_s$ with volume at most p^{-ks}.*

Proof. We shall only prove (i), as (ii) follows from (i) in view of the observation that V_A^M is a translate of V_A^0. It is easy to see that $p^{\rho(x)-s-1} \leqq x < p^{\rho(x)-s}$ for any non-zero $x \in \mathbb{Q}(p^s)$, in view of (2.5) and (2.8).

• Suppose that $\rho(X) \leqq (n-k)s$. For every $j = 1, \ldots, n$, we clearly have $x_j < p^{-a_j}$, where $a_j = s - \rho(x_j)$. Then $a_1 + \cdots + a_n = ns - \rho(X) \geqq ks$, and so Δ_A^0 has volume at most p^{-ks}. Clearly $0 \leqq a_j \leqq s$ for every $j = 1, \ldots, n$, so that $\Delta_A^0 \in \mathfrak{E}_s$.

• Suppose that $\rho(X) > (n-k)s$. For every $j = 1, \ldots, n$, we have $x_j \geqq p^{\rho(x_j)-s-1}$ if $x_j \neq 0$. If $X \in \Delta_A^0$ for some elementary box $\Delta_A^0 \in \mathfrak{E}_s$, then $0 \leqq a_j \leqq s - \rho(x_j)$ for every $j = 1, \ldots, n$, noting that $\rho(0) = 0$. Hence $a_1 + \cdots + a_n \leqq ns - \rho(X) < ks$, and so the volume of Δ_A^0 must exceed p^{-ks}. □

Lemma 7C. *Suppose that a set $D \subseteqq \mathbb{Q}^n(p^s)$ contains exactly p^{ks} points, where $0 \leqq k \leqq n$. Then the following statements are equivalent*:

(i) *D is an optimum $[n, k, s]$-distribution.*

(ii) *The non-Hamming weight $\rho(D) = (n-k)s + 1$.*

Proof. Suppose that D is an optimum $[n, k, s]$-distribution. Then any elementary box in $\mathfrak{E}_s$ with volume p^{-ks} contains exactly one point of D. It follows from Lemma 7B(ii) that $\rho(D) \geqq (n-k)s + 1$. Equality follows in view of Lemma 7A.

Suppose now that $\rho(D) = (n-k)s + 1$. Then it follows from Lemma 7B(ii) that any elementary box in $\mathfrak{E}_s$ with volume p^{-ks} contains at most one point of D. Since D contains exactly p^{ks} points, a density argument now shows that any such elementary box must therefore contain exactly one point of D. □

For any $X, Y \in \mathbb{Q}^n(p^s)$, let $\Phi(X, Y) = e_p(\langle X, Y \rangle)$, where the inner product $\langle X, Y \rangle$ is defined by (2.7), and where $e_p(z) = \mathrm{e}^{2\pi \mathrm{i} z/p}$ for every real number z.

For any function $f\colon \mathbb{Q}^n(p^s) \to \mathbb{C}$, we consider the Fourier transform

$$\tilde{f}(Y) = \sum_{X \in \mathbb{Q}^n(p^s)} \Phi(Y, X) f(X). \tag{7.4}$$

The following result is well known in the theory of abelian groups. For details, see Chapters 5 and 9 of [13] or Chapter 5 of [14].

Lemma 7D. *Suppose that $D, D^\perp \subseteqq \mathbb{Q}^n(p^s)$ are mutually dual linear distributions. Then*

$$\sum_{X \in D} \Phi(Y, X) = \begin{cases} \#(D) & \text{if } Y \in D^\perp, \\ 0 & \text{if } Y \notin D^\perp. \end{cases} \tag{7.5}$$

Furthermore, for any function $f\colon \mathbb{Q}^n(p^s) \to \mathbb{C}$, we have the Poisson summation formula

$$\sum_{X \in D} f(X) = p^{-ns} \#(D) \sum_{Y \in D^\perp} \tilde{f}(Y). \tag{7.6}$$

Suppose that $A = (a_1, \dots, a_n) \in \mathbb{N}_0^n$ satisfies $0 \leqq a_j \leqq s$ for every $j = 1, \dots, n$. For convenience, we shall write $A^* = (a_1^*, \dots, a_n^*) \in \mathbb{N}_0^n$, where $a_j + a_j^* = s$ for every $j = 1, \dots, n$.

Lemma 7E. *Suppose that $A = (a_1, \dots, a_n) \in \mathbb{N}_0^n$ satisfies $0 \leqq a_j \leqq s$ for every $j = 1, \dots, n$. Then*

(i) $(V_A^0)^\perp = V_{A^*}^0$; *and*

(ii) *for every $Y \in \mathbb{Q}^n(p^s)$, we have $\tilde{\chi}_A^0(Y) = p^{ns - a_1 - \cdots - a_n} \chi_{A^*}^0(Y)$, where χ_A^0 and $\chi_{A^*}^0$ denote respectively the characteristic functions of the sets V_A^0 and $V_{A^*}^0$.*

Proof. It is easy to see that V_A^0 consists of points $X = (x_1, \dots, x_n) \in \mathbb{Q}^n(p^s)$ with coordinates

$$x_j = \sum_{i=1}^{s-a_j} \xi_i(x_j) p^{i-s-1}$$

for every $j = 1, \dots, n$, so that $(V_A^0)^\perp$ consists of points $Y = (y_1, \dots, y_n) \in \mathbb{Q}^n(p^s)$ which satisfy the equation $\xi_{s+1-i}(y_j) = 0$ for every $j = 1, \dots, n$ and $i = 1, \dots, s - a_j$; or in equivalent form, $\xi_i(y_j) = 0$ for every $j = 1, \dots, n$ and $i \in [s + 1 - a_j^*, s]$. This gives (i). On the other hand, (ii) is a simple consequence of (7.4), (7.5) and the observation that $\#(V_A^0) = p^{ns - a_1 - \cdots - a_n}$. $\square$

Lemma 7F. *Suppose that $D, D^\perp \subseteqq \mathbb{Q}^n(p^s)$ are mutually dual linear distributions. Then for any elementary box $\Delta_A^M \in \mathfrak{E}_s$, we have*

$$\#(D \cap \Delta_A^M) = p^{-a_1 - \cdots - a_n} \#(D) \sum_{Y \in D^\perp} \Phi(Y, Y_A^M) \chi_{A^*}^0(Y).$$

Proof. Let χ_A^M denote the characteristic function of the set V_A^M. Then it follows from the Poisson summation formula (7.6) that

$$\#(D \cap \Delta_A^M) = \sum_{X \in D} \chi_A^M(X) = p^{-ns} \#(D) \sum_{Y \in D^\perp} \tilde{\chi}_A^M(Y). \tag{7.7}$$

On the other hand, it follows from (7.3) that $\chi_A^M(X) = \chi_A^0(X \ominus Y_A^M)$, and so, in view of (7.4), we have

$$\tilde{\chi}_A^M(Y) = \sum_{X \in \mathbb{Q}^n(p^s)} \Phi(Y,X)\chi_A^0(X \ominus Y_A^M) = \sum_{X \in \mathbb{Q}^n(p^s)} \Phi(Y, X \oplus Y_A^M)\chi_A^0(X) \tag{7.8}$$

$$= \Phi(Y, Y_A^M) \sum_{X \in \mathbb{Q}^n(p^s)} \Phi(Y,X)\chi_A^0(X) = \Phi(Y, Y_A^M)\tilde{\chi}_A^0(Y).$$

The result now follows on combining (7.7), (7.8) and Lemma 7E(ii). □

Note that in the special case $M = 0$ of Lemma 7F, we can take $Y_A^0 = 0$. Since $\Phi(Y, 0) = 1$ for every $Y \in D^\perp$, Lemma 2A follows immediately.

We next need a simple result in the same spirit as Lemma 7B. It can be established by an argument very much similar to the proof of Lemma 7B, so we omit the proof. See Lemma 3.3 of [26].

Lemma 7G. *Let* $\mathscr{B}(t) = \{X \in \mathbb{Q}^n(p^s) : \rho(X) \leqq t\}$. *Then*

$$\mathscr{B}(t) = \bigcup_{a_1 + \cdots + a_n \leqq t} \Delta_{A^*}^0,$$

where every elementary box in the union belongs to $\mathfrak{E}_s$.

Lemma 7H. *Suppose that* $D, D^\perp \leqq \mathbb{Q}^n(p^s)$ *are mutually dual linear distributions of dimensions* d *and* $ns - d$ *respectively. Then for any integer* δ *satisfying* $0 \leqq \delta \leqq d$, *the following statements are equivalent*:

(i) *Each elementary box in* $\mathfrak{E}_s$ *of volume* $p^{\delta - d}$ *contains exactly* p^δ *points of* D.

(ii) *The non-Hamming weight* $\rho(D^\perp) \geqq d + 1 - \delta$.

Proof. Suppose that (i) holds. Then by Lemma 2A, we have $\#(D^\perp \cap \Delta_{A^*}^0) = 1$ whenever $a_1^* + \cdots + a_n^* = ns - d + \delta$. It follows from Lemma 7G that the ball of radius $d - \delta$ contains no point of $D^\perp$ apart from the point 0, and so $\rho(D^\perp) \geqq d + 1 - \delta$. On the other hand, suppose that (ii) holds. Then by Lemma 7G, the set $D^\perp \cap \Delta_{A^*}^0$ contains only the point 0 whenever $a_1^* + \cdots + a_n^* = ns - d + \delta$. It follows from Lemma 7F that for any elementary box Δ_A^M in $\mathfrak{E}_s$ of volume $p^{\delta - d}$, we have

$$\#(D \cap \Delta_A^M) = p^\delta \Phi(0, Y_A^M) = p^\delta. \quad \square$$

Note that the special case $d = s$ gives Lemma 2C. Also, we shall need the following consequence in the next section.

Lemma 7I. *A subset* $D \leqq \mathbb{Q}^n(p^s)$ *is a linear optimum* $[n, k, s]$*-distribution if and only if its dual* $D^\perp$ *is a linear optimum* $[n, n - k, s]$*-distribution.*

Proof. Using Lemma 7H with $d = ks$ and $\delta = 0$, we deduce that D is a linear optimum $[n, k, s]$-distribution if and only if $\rho(D^\perp) \geqq ks + 1$, if and only if $\rho(D^\perp) = ks + 1$ in view of Lemma 7A. On the other hand, it follows from Lemma 7C that $D^\perp$ is a linear optimum $[n, n - k, s]$-distribution if and only if $\rho(D^\perp) = ks + 1$. □

8. Isomorphisms of vector spaces $\mathbb{Q}^{gn}(p^\sigma)$ and $\mathbb{Q}^n(p^{g\sigma})$

Suppose that $\sigma, g \in \mathbb{N}$. The mapping $\pi\colon \mathbb{Q}^g(p^\sigma) \to \mathbb{Q}(p^{g\sigma})$, where

$$\pi(\omega_1, \dots, \omega_g) = \sum_{\ell=1}^{g} p^{-(\ell-1)\sigma}\omega_\ell$$

for every $(\omega_1, \dots, \omega_g) \in \mathbb{Q}^g(p^\sigma)$, is related to the well known Peano mapping which gives a bijection between points in U^g and points in U, restricted here to $\mathbb{Q}^g(p^\sigma)$. It is easy to see that π gives an isomorphism between the vector spaces $\mathbb{Q}^g(p^\sigma)$ and $\mathbb{Q}(p^{g\sigma})$.

We now extend π to a mapping $\Pi\colon \mathbb{Q}^{gn}(p^\sigma) \to \mathbb{Q}^n(p^{g\sigma})$ by writing

$$\Pi(\varpi_1, \dots, \varpi_n) = \big(\pi(\varpi_1), \dots, \pi(\varpi_n)\big)$$

for every $\varpi_1, \dots, \varpi_n \in \mathbb{Q}^g(p^\sigma)$. It is easy to see that Π gives an isomorphism between the vector spaces $\mathbb{Q}^{gn}(p^\sigma)$ and $\mathbb{Q}^n(p^{g\sigma})$. The following result is a simple consequence of this observation.

Lemma 8A. *Suppose that $D \subseteqq \mathbb{Q}^{gn}(p^\sigma)$. Then*

(i) $\#(D) = \#\big(\Pi(D)\big)$; *and*

(ii) *if D is a linear distribution, then $\Pi(D)$ is also a linear distribution.*

We next study the effect of the mapping Π on the metrics $\varkappa$ and ρ.

Lemma 8B. (i) *For every $\Omega \in \mathbb{Q}^{gn}(p^\sigma)$, we have*

$$\varkappa\big(\Pi(\Omega)\big) = \varkappa(\Omega) \quad \textit{and} \quad \rho\big(\Pi(\Omega)\big) \geqq \rho(\Omega).$$

(ii) *For every $D \subseteqq \mathbb{Q}^{gn}(p^\sigma)$, we have $\varkappa\big(\Pi(D)\big) = \varkappa(D)$ and $\rho\big(\Pi(D)\big) \geqq \rho(D)$.*

Proof. The result for $\varkappa$ follows easily as the number of non-zero coefficients remains unchanged. Suppose now that $(\omega_1, \dots, \omega_g) \in \mathbb{Q}^g(p^\sigma)$ satisfies $\omega_1 = \dots = \omega_{\ell-1} = 0$ and $\omega_\ell \neq 0$. Then $\rho(\omega_1) = \dots = \rho(\omega_{\ell-1}) = 0$, and

$$\rho\big(\pi(\omega_1, \dots, \omega_g)\big) = (g-\ell)\sigma + \rho(\omega_\ell) \geqq \rho(\omega_g) + \dots + \rho(\omega_\ell) = \rho(\omega_1, \dots, \omega_g).$$

The assertions for ρ now follow easily. □

Remark. The validity of Lemma 8B for the weight ρ dictates our choice for the isomorphisms π and Π. It should be mentioned that there are other choices for such isomorphisms. Indeed, in the paper [26], such isomorphisms were defined in terms of the well known Peano mapping which gives a bijection between points in the unit cubes U^{gn} and U^n, restricted to $\mathbb{Q}^{gn}(p^\sigma)$ and $\mathbb{Q}^n(p^{g\sigma})$ respectively. However, our present choice of the isomorphisms π and Π makes the proof of Lemma 8B a little bit simpler.

Lemma 8C. *Suppose that $D \subseteqq \mathbb{Q}^{gn}(p^\sigma)$ is an optimum $[gn, gk, \sigma]$-distribution. Then $\rho\big(\Pi(D)\big) = (n-k)g\sigma + 1$ and $\varkappa\big(\Pi(D)\big) \geqq (n-k)g + 1$.*

Proof. In view of Lemma 7C, we have $\rho(D) = (n-k)g\sigma + 1$. It follows from Lemma 8B that $\rho\big(\Pi(D)\big) \geqq (n-k)g\sigma + 1$. Equality follows in view of Lemma 7A. Next, note that for every $D \subseteqq \mathbb{Q}^{gn}(p^\sigma)$, we have $\varkappa(D) \geqq \sigma^{-1}\rho(D) = (n-k)g + \sigma^{-1}$, so that $\varkappa(D) \geqq (n-k)g + 1$. The second assertion now follows from Lemma 8B. $\square$

To study dual distributions, we need to introduce a reflection mapping.

For every $\varpi = (\omega_1, \dots, \omega_g) \in \mathbb{Q}^g(p^\sigma)$, we consider the reflection

$$\varpi^{\leftrightarrow} = (\omega_g, \dots, \omega_1) \in \mathbb{Q}^g(p^\sigma).$$

It is easy to check that for every $\varpi_1, \varpi_2 \in \mathbb{Q}^g(p^\sigma)$, we have

$$\langle \pi(\varpi_1), \pi(\varpi_2) \rangle = \langle \varpi_1^{\leftrightarrow}, \varpi_2 \rangle = \langle \varpi_1, \varpi_2^{\leftrightarrow} \rangle. \tag{8.1}$$

For every $\Omega = (\varpi_1, \dots, \varpi_n) \in \mathbb{Q}^{gn}(p^\sigma)$, we now let

$$\Omega^{\leftrightarrow} = (\varpi_1^{\leftrightarrow}, \dots, \varpi_n^{\leftrightarrow}) \in \mathbb{Q}^{gn}(p^\sigma).$$

It follows from (8.1) that for every $\Omega_1, \Omega_2 \in \mathbb{Q}^{gn}(p^\sigma)$, we have

$$\langle \Pi(\Omega_1), \Pi(\Omega_2) \rangle = \langle \Omega_1^{\leftrightarrow}, \Omega_2 \rangle = \langle \Omega_1, \Omega_2^{\leftrightarrow} \rangle. \tag{8.2}$$

For any $D \subseteqq \mathbb{Q}^{gn}(p^\sigma)$, write $D^{\leftrightarrow} = \{\Omega^{\leftrightarrow} \colon \Omega \in D\}$.

Lemma 8D. *Suppose that $D \subseteqq \mathbb{Q}^{gn}(p^\sigma)$ is a linear distribution. Then*

$$\big(\Pi(D)\big)^{\perp} = \Pi\big((D^{\perp})^{\leftrightarrow}\big).$$

Proof. We can write $\big(\Pi(D)\big)^{\perp} = \Pi(D_1)$. Then it follows from (8.2) that $D_1^{\leftrightarrow} = D^{\perp}$, so that $D_1 = (D^{\perp})^{\leftrightarrow}$. $\square$

Lemma 8E. *Suppose that $D \subseteqq \mathbb{Q}^{gn}(p^\sigma)$ is a linear optimum $[gn, gk, \sigma]$-distribution. Then $\rho\big(\big(\Pi(D)\big)^{\perp}\big) \geqq kg\sigma + 1$ and $\varkappa\big(\big(\Pi(D)\big)^{\perp}\big) \geqq kg + 1$.*

Proof. Note first of all that for every $D \subseteqq \mathbb{Q}^{gn}(p^\sigma)$, we have

$$\rho(D^{\leftrightarrow}) = \rho(D) \quad \text{and} \quad \varkappa(D^{\leftrightarrow}) = \varkappa(D). \tag{8.3}$$

Suppose that $D \subseteqq \mathbb{Q}^{gn}(p^\sigma)$ is a linear optimum $[gn, gk, \sigma]$-distribution. Then it follows from Lemma 7I that $D^{\perp}$ is a linear optimum $[gn, g(n-k), \sigma]$-distribution. Hence it follows from Lemma 8D, Lemma 8B, (8.3) and Lemma 7C that

$$\rho\big(\big(\Pi(D)\big)^{\perp}\big) = \rho\big(\Pi\big((D^{\perp})^{\leftrightarrow}\big)\big) \geqq \rho\big((D^{\perp})^{\leftrightarrow}\big) = \rho(D^{\perp}) = kg\sigma + 1.$$

Similarly, we have

$$\varkappa\big(\big(\Pi(D)\big)^{\perp}\big) = \varkappa\big(\Pi\big((D^{\perp})^{\leftrightarrow}\big)\big) = \varkappa\big((D^{\perp})^{\leftrightarrow}\big) = \varkappa(D^{\perp}).$$

Note now that $\varkappa(D^{\perp}) \geqq \sigma^{-1}\rho(D^{\perp}) = kg + \sigma^{-1}$, so that $\varkappa(D^{\perp}) \geqq kg + 1$. $\square$

Let us return to Lemma 2E, and use the notation in Section 2. Consider a set $D \subseteqq \mathbb{Q}^{gn}(p^\sigma)$ of the form

$$D = \big\{\big(\varpi_1(f), \dots, \varpi_n(f)\big) : f \in \mathbb{F}_p[z] \text{ and } \deg f < g\sigma\big\},$$

where for every $i = 1, \dots, n$,

$$\varpi_i(f) = \left(\sum_{j=1}^{\sigma} \partial^{j-1} f(\beta_{i,1}) p^{-j}, \dots, \sum_{j=1}^{\sigma} \partial^{j-1} f(\beta_{i,g}) p^{-j} \right) \in \mathbb{Q}^g(p^\sigma).$$

It is easy to see that D has exactly $p^{g\sigma}$ elements, and that $\Pi(D) = D(g, \sigma)$. It follows from Lemma 8E with $k = 1$ that to prove Lemma 2E, it remains to show that D is a linear optimum $[gn, g, \sigma]$-distribution. Since the collection of polynomials $f \in \mathbb{F}_p[z]$ with $\deg f < g\sigma$ is closed under addition and scalar multiplication in $\mathbb{F}_p$, it follows that D is a linear distribution. Our proof of Lemma 2E is therefore complete if we can establish the following result.

Lemma 8F. *Every elementary box in $\mathfrak{E}_\sigma$ of volume $p^{-g\sigma}$ contains exactly one point of D.*

Proof. Suppose that an elementary box in $\mathfrak{E}_\sigma$ of volume $p^{-g\sigma}$ is chosen. Then the number of points of D that fall into this elementary box is given by the number of solutions of a system

$$\partial^{j-1} f(\beta_{i,\ell}) = a_{i,\ell}^{(j)}, \quad \begin{cases} (i,\ell) \in \mathscr{I} \subseteqq \{1, \dots, n\} \times \{1, \dots, g\}, \\ j = 1, \dots, t_{i,\ell} \text{ where } t_{i,\ell} \leqq \sigma, \\ \sum\limits_{(i,\ell) \in \mathscr{I}} t_{i,\ell} = g\sigma. \end{cases} \tag{8.4}$$

This is the so-called Hermite interpolation problem and has a unique solution $f \in \mathbb{F}_p[z]$ with $\deg f < g\sigma$. To see this, consider the polynomials

$$r_{i,\ell}(z) = \sum_{j=1}^{t_{i,\ell}} a_{i,\ell}^{(j)} (z - \beta_{i,\ell})^{j-1}, \quad (i,\ell) \in \mathscr{I}.$$

The system (8.4) is equivalent to the system of congruences

$$f(z) = r_{i,\ell}(z) \bmod (z - \beta_{i,\ell})^{t_{i,\ell}}, \quad (i,\ell) \in \mathscr{I}. \tag{8.5}$$

Since the polynomials $(z - \beta_{i,\ell})^{t_{i,\ell}}$, where $(i,\ell) \in \mathscr{I}$, are pairwise coprime, it follows that the system (8.5) of congruences has a unique solution $f \in \mathbb{F}_p[z]$ with $\deg f < g\sigma$, in view of the Chinese remainder theorem in the ring $\mathbb{F}_p[z]$. □

References

[1] *J. Beck*, *W. W. L. Chen*, Irregularities of Distribution, Cambridge University Press, Cambridge 1987.

[2] *M. S. Birman*, *M. Z. Solomyak*, Spectral Theory of Self-Adjoint Operators in Hilbert Space, Leningrad State University, Leningrad, 1980; english translation: Reidel, Dordrecht 1987.

[3] *W. W. L. Chen*, On irregularities of distribution, Mathematika **27** (1980), 153–170.
[4] *W. W. L. Chen*, On irregularities of distribution II, Quart. J. Math. Oxford **34** (1983), 257–279.
[5] *W. W. L. Chen*, *M. M. Skriganov*, Davenport's theorem in the theory of irregularities of point distribution, Zapiski Nauch. Sem. POMI **269** (2000), 339–353.
[6] *H. Davenport*, Note on irregularities of distribution, Mathematika **3** (1956), 131–135.
[7] *N. M. Dobrovol'skiĭ*, An effective proof of Roth's theorem on quadratic dispersion, Uspekhi Mat. Nauk **39** (1984), 155–156; english translation: Russian Math. Surv. **39** (1984), 117–118.
[8] *H. Faure*, Discrépance de suites associées à un système de numération (en dimension s), Acta Arith. **41** (1982), 337–351.
[9] *N. J. Fine*, On the Walsh functions, Trans. Amer. Math. Soc. **65** (1949), 373–414.
[10] *K. K. Frolov*, An upper bound for the discrepancy in the L_p-metric, Dokl. Acad. Nauk SSSR **252** (4) (1980), 805–807.
[11] *B. I. Golubov*, *A. V. Efimov*, *V. A. Skvorčov*, The Walsh Series and Transformations—Theory and Applications, Nauka, Moscow 1987; english translation: Kluwer, Dordrecht 1991.
[12] *G. Larcher*, *F. Pillichshammer*, On the L_2-discrepancy of the Sobol-Faure-Niederreiter net in dimension 3, J. Complexity, to appear.
[13] *R. Lidl*, *H. Niederreiter*, Finite fields, Addison-Wesley 1983.
[14] *F. J. MacWilliams*, *N. J. A. Sloane*, The Theory of Error-Correcting Codes, North-Holland, Amsterdam 1977.
[15] *W. J. Martin*, *D. R. Stinson*, Association schemes for ordered orthogonal arrays and (T, M, S)-nets, Canadian J. Math. **51** (1999), 326–346.
[16] *H. Niederreiter*, Point sets and sequences with small discrepancy, Monatsh. Math. **104** (1987), 273–337.
[17] *H. Niederreiter*, Nets, (t, s)-sequences, and algebraic curves over finite fields with many rational points, International Congress of Mathematicians in Berlin, Extra Volume III, Documenta Mathematica (1998), 337–386.
[18] *J. J. Price*, Certain groups of orthonormal step functions, Canadian J. Math. **9** (1957), 413–425.
[19] *M. Yu. Rosenbloom*, *M. A. Tsfasman*, Codes in the m-metric, Problemi Peredachi Inf. **33** (1) (1997), 55–63; english translation: Probl. Inf. Trans. **33** (1) (1997), 45–52.
[20] *K. F. Roth*, On irregularities of distribution, Mathematika **1** (1954), 73–79.
[21] *K. F. Roth*, On irregularities of distribution III, Acta Arith. **35** (1979), 373–384.
[22] *K. F. Roth*, On irregularities of distribution IV, Acta Arith. **37** (1980), 67–75.
[23] *F. Schipp*, *W. R. Wade*, *P. Simon*, Walsh Functions—An Introduction to Dyadic Harmonic Analysis, Adam Hilger, Bristol and New York 1990.
[24] *M. M. Skriganov*, Lattices in algebraic number fields and uniform distributions modulo 1, Algebra i Analiz **1** (2) (1989), 207–228; english translation: Leningrad Math. J. **1** (1990), 535–558.
[25] *M. M. Skriganov*, Constructions of uniform distributions in terms of geometry of numbers, Algebra i Analiz **6** (3) (1994), 200–230; reprinted in St. Petersburg Math. J. **6** (1995), 635–664.
[26] *M. M. Skriganov*, Coding theory and uniform distributions, Algebra i Analiz **13** (2) (2001), 191–239; english translation: St. Petersburg Math. J., to appear.
[27] *I. M. Sobol*, On the distribution of points in a cube and the approximate evaluation of integrals, Ž. Vyčisl. Mat. i Mat. Fiz. **7** (1967), 784–802; english translation: USSR Comp. Math. and Math. Phys. **7** (1967), 86–112.

Department of Mathematics, Macquarie University, Sydney NSW 2109, Australia
e-mail: wchen@math.mq.edu.au

Steklov Mathematical Institute, Fontanka 27, St Petersburg 191011, Russia
e-mail: skrig@pdmi.ras.ru

Eingegangen 27. September 2000