SOME DETAILS OF THE ELLENBERG-GIJSWIJT PROOF OF THE CAP SET PROBLEM

1. Introduction

In order to formalize, e.g. in the theorem prover *Lean*, the Ellenberg-Gijswijt proof [1] of the *cap set problem* (and generalization) we spell out most details of this proof, basically starting at Section 10 (and a bit in the two sections before). For the asymptotics at the end, we follow a different method. Furthermore, we write out several details of the preliminary underlying mathematics (mainly linear algebra).

NOTE: the only actual intended use of this informal document was for Johannes, Rob, and Sander to work from in combination with discussions, explanations, etc. In particular, with very few exceptions, no attempt was made to correct errors or complete omissions unless still beneficial for the formalization process.

2. Notation and conventions

Throughout this whole article, k denotes a field.

A *ring* R contains a one, denoted 1_R (which could be equal to zero, denoted 0_R). A *ring homomorphism* $R \to S$ sends 1_R to 1_S.

When we write 'R-module' in a statement, then this can be read throughout the whole statement as 'left R-module' or throughout the whole statement as 'right R-module'.

We use '$\subset$' for inclusion, and '$\subsetneq$' for strict inclusion.

3. Very basic lemmata for sets

Lemma 3.1. *Let A and B be sets. Then the following hold.*

(i) $|A \cup B| + |A \cap B| = |A| + |B|$.

(ii) $|A \cup B| \leq |A| + |B|$.

Lemma 3.2. *Let A and B be sets with $A \subsetneq B$ and A finite. Then $|A| < |B|$.*

4. Very basic lemmata for groups and fields

Lemma 4.1. *Let x be an element of finite order m in a group (with identity e), and $n \in \mathbb{Z}$. Then*

$$m|n \Leftrightarrow x^n = e.$$

Proof. Note $m \in \mathbb{Z}_{>0}$. By devision with remainder we have unique $q, r \in \mathbb{Z}$ with $n = qm + r$ and $0 \leq r < m$. Note

$$m|n \Leftrightarrow r = 0.$$

Since $0 \leq r < m = \mathrm{ord}(x)$ we get

$$r = 0 \Leftrightarrow x^r = e.$$

Using $n = qm + r$ and $x^m = e$ gives: $x^n = (x^m)^q x^r = e^q x^r = x^r$, so

$$x^r = e \Leftrightarrow x^n = e.$$

Composing the three equivalences above, yields the desired result. □

Here is a quick corollary to Lagrange's theorem in group theory

Lemma 4.2. *Let G be a finite group (with identity e) and $x \in G$. Then $x^{|G|} = e$.*

Proof. We give a quick proof in the case that G is abelian (which is the only case we need). The general case follows of course quickly from Lagrange's theorem. Note that $G \to G : g \mapsto xg$ is bijective. Now assume G is abelian. We get that $\prod_{g\in G} g = \prod_{g\in G} xg = x^{|G|} \prod_{g\in G} g$. The result follows from multiplying the left and right hand side by $\left(\prod_{g\in G} g\right)^{-1}$. □

Lemma 4.3. *Let $\mathbb{F}_q$ be a finite field with q elements and $x \in \mathbb{F}_q$. Then*

$$x^{q-1} = \begin{cases} 0, & \text{if } x = 0; \\ 1, & \text{if } x \neq 0. \end{cases}$$

Proof. If $x = 0$, then obviously $x^{q-1} = 0^{q-1} = 0$ (since $q > 1$). Now let $x \neq 0$. Then $x \in \mathbb{F}_q^*$, which is a group of order $q - 1$ with identity 1. So by Lemma 4.2 we get $x^{q-1} = 1$. □

The previous lemma immediately translates to the following.

Lemma 4.4. *Let $\mathbb{F}_q$ be a finite field with q elements and $x, a \in \mathbb{F}_q$. Then*

$$1 - (x - a)^{q-1} = \begin{cases} 1, & \text{if } x = a; \\ 0, & \text{if } x \neq a. \end{cases}$$

5. Very basic lemmata for modules

Throughout this section, R denotes a ring.

Kernel and image of a module are modules.

Lemma 5.1. *Let M and N be R-modules and $f \in \mathrm{Hom}(V, W)$.*

(i) $\ker(f)$ *is an R-submodule of M.*
(ii) $\mathrm{im}(f)$ *is an R-submodule of N.*

Proof. Just write out definitions ... □

The restriction of a morphism to a submode is a morphism and the corresponding image is a submodule.

Lemma 5.2. *Let M and N be R-modules, P an R-submodule of M, and $f \in \mathrm{Hom}(M, N)$. Then the following hold.*

(i) $f|_P$ (the restriction of f to P) is in $\mathrm{Hom}(P, N)$.
(ii) $f(P)$ is an R-submodule of N.

Proof. 'i': Just write out definitions ...

'ii': $f|_P$ is in $\mathrm{Hom}(P, N)$ by part i. So $\mathrm{im}(f|_P)$ is an R-submodule of N by Lemma 5.1 part (ii). Finally, note $f(P) = f|_P(P) = \mathrm{im}(f|_P)$. □

6. Some basics for vector spaces

Lemma 6.1. *Let V and W be finite dimensional k-vector spaces with $\dim(V) = \dim(W)$. Let $f \in \mathrm{Hom}(V, W)$. Then the following statements are equivalent.*

(i) *f is injectve.*
(ii) *f is surjective.*
(iii) *f is bijective.*
(iv) *$f \in \mathrm{Isom}(V, W)$.*

For the part above Propostion 2 in Elleberg-Gijswijt, we actually only need '(ii) $\Rightarrow$ (iv)'.

A spanning set contains a basis (accepting the axiom of choice).

Lemma 6.2. *Let V be a k-vector space and S a subset of V that spans V. Then S contains a basis of V.*

Proof. Consider the collection of all linearly independent subsets of S, ordered by inclusion. By Zorn's lemma this collection contains a maximal element $B \subset S$. One checks $(\dots)$ that such B must be a basis of V. □

A linearly independent set can be enlarged to a basis (accepting the axiom of choice).

Lemma 6.3. *Let V be a k-vector space and S a linearly independent subset of V. Then S is contained in a basis of V.*

Proof. Consider the collection of all linearly independent supersets of S, ordered by inclusion. By Zorn's lemma this collection contains a maximal element $B \subset S$. One checks $(\dots)$ that such B must be a basis of V. □

Remark 6.4. If V is finite dimensional, we do of course not need the axiom of choice.

The Rank-nullity theorem.

Theorem 6.5. *Let V and W be k-vector spaces and $f \in \mathrm{Hom}(V, W)$. Then*

$$\dim(\mathrm{im}(f)) + \dim(\ker(f)) = \dim(V).$$

Lemma 6.6. *Let V be a k-vector space and X a linear subspace of V. Then the following hold.*

(i) *$\dim(X) \leq \dim(V)$.*
(ii) *If $\dim(X)$ is finite and $X \neq V$, then $\dim(X) < \dim(V)$.*

Proof. Let B be basis for X. By Lemma 6.3 we have a basis B' for V with $B \subset B'$. Part (i) follows immediately.

For part (ii), assume that $\dim(X)$ is finite and $X \neq V$. So B is finite and $B \subsetneq B'$. By Lemma 3.2 we get that $|B| < |B'|$. This means by definition that $\dim(X) < \dim(V)$. □

Lemma 6.7. *Let V and W be k-vector spaces, X be a linear subspace of V, and $f \in \mathrm{Hom}(V, W)$. Then the following hold.*

(i) *$f(X)$ is a linear subspace of W.*
(ii) *$\dim(f(X)) \leq \dim(V)$.*
(iii) *$\dim(f(X)) \leq \dim(W)$.*
(iv) *$\dim(f(X)) \leq \dim(X)$.*

Proof. 'i': Specialize $R = k$ in Lemma 5.2 part (ii).
Rest: also very basic (several possibilities...). □

Definition 6.8. Let V be a k-vector space and W_1, W_2 linear subspaces of V. Then

$$W_1 + W_2 := \{w_1 + w_2 \in V : w_1 \in W_1 \text{ and } w_2 \in W_2\}. \tag{1}$$

Lemma 6.9. *Let V be a k-vector space and W_1, W_2 linear subspaces of V. Then $W_1 + W_2$ is a linear subspace of V.*

Proof. Write out definitions. □

Lemma 6.10. *Let V be a k-vector space and W_1, W_2 linear subspaces of V. Then*

$$\dim(W_1 + W_2) \leq \dim(W_1) + \dim(W_2).$$

Proof. Let B_1, B_2 be bases for W_1, W_2 respectively. Then $B_1 \cup B_2$ spans $W_1 + W_2$ (write out definitions). By Lemma 6.2 we have a subset $S \subset B_1 \cup B_2$ such that S is a basis for $W_1 + W_2$. Now

$$\begin{aligned}
\dim(W_1 + W_2) &= |S| && \text{(by definition since } S \text{ is a basis for } W_1 + W_2) \\
&\leq |B_1 \cup B_2| && \text{(since } S \subset B_1 \cup B_2) \\
&\leq |B_1| + |B_2| && \text{(by Lemma 3.1 part (ii))} \\
&= \dim(W_1) + \dim(W_2) && \text{(by definition).}
\end{aligned}$$

□

7. Ranks

Definition 7.1. Let V and W be k-vector spaces and $f \in \mathrm{Hom}(V, W)$. Then

$$\mathrm{rank}(f) := \dim(\mathrm{im}(f)).$$

It is basically immediate that the rank is bounded by the dimensions of the domain and codomain.

Lemma 7.2. *Let V and W be k-vector spaces and $f \in \mathrm{Hom}(V, W)$. Then*

$$\mathrm{rank}(f) \leq \min(\dim(V), \dim(W)).$$

Proof. Apply e.g. Lemma 6.7. □

An important basic property is that the rank is subadditive.

Proposition 7.3. *Let V and W be k-vector spaces and $f, g \in \mathrm{Hom}(V, W)$. Then $\mathrm{rank}(f + g) \leq \mathrm{rank}(f) + \mathrm{rank}(g)$.*

Proof. We claim that

$$\mathrm{im}(f + g) \subset \mathrm{im}(f) + \mathrm{im}(g). \tag{2}$$

Indeed, let $w \in \mathrm{im}(f + g)$. Then we have $v \in V$ such that $w = (f + g)(v) = f(v) + g(v) \in \mathrm{im}(f) + \mathrm{im}(g)$, which proves the claim. Now

$$\begin{aligned}
\mathrm{rank}(f + g) &= \dim(\mathrm{im}(f + g)) && \text{(by definition)} \\
&\leq \dim(\mathrm{im}(f) + \mathrm{im}(g)) && \text{(by (2) and Lemma 6.6 part (i))} \\
&\leq \dim(\mathrm{im}(f)) + \dim(\mathrm{im}(g)) && \text{(by Lemma 6.10)} \\
&= \mathrm{rank}(f) + \mathrm{rank}(g) && \text{(by definition).}
\end{aligned}$$

□

Next, a rank-property for multiplying (i.e. composing) linear maps.

Proposition 7.4. *Let U, V, and W be k-vector spaces and $f \in \mathrm{Hom}(V,W)$, $g \in \mathrm{Hom}(U,V)$. Then*

$$\mathrm{rank}(fg) \leq \min(\mathrm{rank}(f), \mathrm{rank}(g)).$$

Proof. Note that is suffices to prove

$$\mathrm{rank}(fg) \leq \mathrm{rank}(f) \tag{3}$$

and to prove

$$\mathrm{rank}(fg) \leq \mathrm{rank}(g). \tag{4}$$

We start with proving (3). $g(U) \subset V$, hence $f(g(U)) \subset f(V)$, hence (using Lemma 6.6 part (i)) $\dim(f(g(U))) \leq \dim(f(V))$, which translates by definition to $\mathrm{rank}(fg) \leq \mathrm{rank}(f)$.

Next, we prove (4). By Lemma 6.7 part (iv) with $X = g(U)$ (which is a linear subspace by the same Lemma part (i)) we get $\dim(f(g(U))) \leq \dim(g(U))$, which translates by definition to $\mathrm{rank}(fg) \leq \mathrm{rank}(g)$. □

8. Matrices

For finite sets A, B and vector spaces $V := k^A, W := k^B$ we identify $f \in \mathrm{Hom}(V,W)$ with the $B \times A$ matrix associated to bases of indicator functions of the singletons for A and B respectively.

Below are rank bounds for matrices of a specific form.

Lemma 8.1. *Let A be a finite set, $x, y : A \to k$ functions, and consider the $A \times A$ matrix M with entry $M_{a,b} := x(a)y(b)$ (for $a, b \in A$). Then* $\mathrm{rank}(M) \leq 1$.

Proof. Let $\{\bullet\}$ denote a singelton. Consider the linear maps $X : k^{\{\bullet\}} \to k^A$ with matrix $X_{a,\bullet} := x(a)$ (for $a \in A$) and $Y : k^A \to k^{\{\bullet\}}$ with matrix $Y_{\bullet,a} := y(a)$ (for $a \in A$). One immediately checks that we have the factorization $M = XY$. So by Proposition 7.4 we get $\mathrm{rank}(M) \leq \min(\mathrm{rank}(X), \mathrm{rank}(Y))$. By Lemma 7.2 we have $\mathrm{rank}(X) \leq 1$ and $\mathrm{rank}(Y) \leq 1$. We conclude that $\mathrm{rank}(M) \leq 1$. □

Lemma 8.2. *Let A be a finite set and M an $A \times A$ diagonal matrix, i.e. $M_{a,b} = 0$ if $a \neq b$ (for $a, b \in A$). Then*

$$\mathrm{rank}(M) = |\{a \in A : M_{a,a} \neq 0\}|.$$

Proof. Compute $\mathrm{im}(M) = \mathrm{span}\{\, \iota_a \mid a \in A, M_{a,a} \neq 0 \,\}$ where ι_a denotes the indicator function of $\{a\}$ ... □

Perhaps it is easier to write everything in terms of $n \times n$ matrices.

Lemma 8.3. *Let $n \in \mathbb{Z}_{\geq 0}$ and $x := (x_1, \ldots, x_n), y := (y_1, \ldots, y_n) \in k^n$. Consider the $n \times n$ matrix M with entry $M_{i,j} := x_i y_j$ (for $1 \leq i, j \leq n$). Then* $\mathrm{rank}(M) \leq 1$.

Proof. By definition of matrix multiplication we observe that $M = x^t y$ (so x^t is a column vector and y a row vector). So $\mathrm{rank}(M) \leq \min(\mathrm{rank}(x^t), \mathrm{rank}(y)) \leq 1$. □

Lemma 8.4. *Let $n \in \mathbb{Z}_{\geq 0}$ and M an $n \times n$ diagonal matrix, i.e. $M_{i,j} = 0$ if $i \neq j$ (for $1 \leq i, j \leq n$). Then*

$$\mathrm{rank}(M) = |\{\, i \in \{1, \ldots, n\} \mid M_{i,i} \neq 0 \,\}|.$$

Proof. Just write out $\mathrm{im}(M)$... □

9. Further preliminaries

For R a ring and $n \in \mathbb{Z}_{\geq 0}$, the polynomial ring $R[x_1, \ldots, x_n]$ has a canonical structure as an R-module, which is free with basis all monomials $M_B := \{\prod_{i=1}^n x_i^{e_i} : e_1, \ldots, e_n \in \mathbb{Z}_{\geq 0}\}$. (Note that when $n = 0$ we simply get R back with $M_B = \{1\}$.) Now in case $R = k$ (a field) we get a k-vectorspace, whose dimension is (countably) infinite when $n \neq 0$. Of course, any subset $S \subset M_B$ is the basis of the linear subspace $\mathrm{span}_k(S)$ (of dimension $|S|$).

For every $a \in R^n$ we have the (local) evaluation homomorphism (say of R-modules) $\mathrm{ev}_a : R[x_1, \ldots, x_n] \to R$ given by $f \mapsto f(a)$. This yields a (global) evaluation homomorphism $\mathrm{ev} : R[x_1, \ldots, x_n] \to (R^n \to R)$ given by $f \mapsto (a \mapsto \mathrm{ev}_a(f))$, i.e. simply the canonical function associated to a polynomial. Note that $R^n \to R = R^{R^n}$ consists of all function from R^n to R and has (canonically) the structure of an R-module.

Lemma 9.1. *Let $\mathbb{F}_q, n, M_n$, and S_n be as defined in Section 10. Then the restriction (of $\mathrm{ev} : \mathbb{F}_q[x_1, \ldots, x_n] \to \mathbb{F}_q^{\mathbb{F}_q^n}$ to S_n) $\mathrm{ev}' := \mathrm{ev}|_{S_n} : S_n \to \mathbb{F}_q^{\mathbb{F}_q^n}$ is an isomorphism of k-vectorspaces.*

Proof. Note that S_n is a linear subspace of dimension $|M_n| = q^n$ of the k-vectorspace $k[x_1, \ldots, x_n]$ since S_n is by definition the span of M_n (a subset of the standard monomial basis of $k[x_1, \ldots, x_n]$). Write $A := \mathbb{F}_q^n$, then $\mathbb{F}_q^{\mathbb{F}_q^n} = \mathbb{F}_q^A$ has dimension $|A| = q^n$ (as $\mathbb{F}_q$-vectorspace). So $\dim(S_n) = \dim(\mathbb{F}_q^A)$ and bij Lemma 6.1 it suffices to prove that ev' is surjective. Since the indicator functions $\iota_a : A \to \mathbb{F}_q$ (given by $\iota_a(x) = 1$ if $x = a$ and $\iota_a(x) = 0$ otherwise) for $a \in A$ give a basis for $\mathbb{F}_q^A$, it suffices that these are in the image of ev'. To show the latter, let $a \in A$ and define $f := \prod_{i=1}^n \left(1 - (x_i - a_i)^{q-1}\right) \in S_n$. Then by Lemma 4.4 we readily see that $\mathrm{ev}'(f) = \iota_a$. □

Lemma 9.2. *Let V be a linear subspace of the k-vectorspace $k[x_1, \ldots, x_n]$, let $A \subset k^n$, and let*

$$X := \{p \in V : \forall a \in A : p(a) = 0\}.$$

Then X is a linear subspace of V with

$$\dim(X) + |A| \geq \dim(V).$$

Proof. Consider the map

$$f : V \to k^A : \quad p \mapsto (p(a))_{a \in A}.$$

Note that this is a linear map with

$$\ker(f) = X$$

and

$$\dim(\mathrm{im}(f)) \leq \dim(k^A) = |A|.$$

By Theorem 6.5

$$\dim(V) = \dim(X) + \dim(\mathrm{im}(f)) \leq \dim(X) + |A|.$$

The lemma follows. □

10. Notation from [1]

- $\mathbb{F}_q$: finite field with q elements
- $n \in \mathbb{Z}_{>0}$
- $\mathbb{F}_q[x_1, \dots, x_n]$: the ring of polynomials in n variables with coefficients in $\mathbb{F}_q$; Note that it also has the canonical structure of an $\mathbb{F}_q$-vector space (of infinite dimension).
- $M_n := \{\prod_{i=1}^n x_i^{a_i} \in \mathbb{F}_q[x_1, \dots, x_n] : 0 \le a_i \le q-1\}$; note $|M_n| = q^n$.
- $S_n := \{\sum_{m \in M_n} c_m m : c_m \in \mathbb{F}_q\}$: the $\mathbb{F}_q$-vector space spanned by M_n; note that M_n is a basis, hence $\dim(S_n) = |M_n| = q^n$.
- For $d \in [0, (q-1)n]$ (the latter is an interval in the real numbers)

  $$M_n^d := \{m \in M_n : \deg(m) \le d\}$$
  $$= \{\prod_{i=1}^n x_i^{a_i} \in \mathbb{F}_q[x_1, \dots, x_n] : 0 \le a_i \le q-1 \quad \text{and} \quad \sum_{i=1}^n a_i \le d\}.$$

- $S_n^d := \{\sum_{m \in M_n^d} c_m m : c_m \in \mathbb{F}_q\}$: the $\mathbb{F}_q$-vector space spanned by M_n^d; note that M_n^d is a basis.
- $m_d := \dim(S_n^d) = |M_n^d|$.

11. Proof of Proposition 2 from [1]

Proposition 11.1 (Proposition 2 from [1])**.** *Let $A \subset \mathbb{F}_q^n$, $\alpha, \beta, \gamma \in \mathbb{F}_q$ such that*

$$\alpha + \beta + \gamma = 0, \tag{5}$$

and $P \in S_n^d$ such that for all $a, b \in A$ with $a \ne b$:

$$P(\alpha a + \beta b) = 0. \tag{6}$$

Then

$$|\{a \in A : P(-\gamma a) \ne 0\}| \le 2m_{d/2}.$$

Let $\alpha, \beta, \gamma, A, P$ be as in the assumptions..

⋮

See the first 9 lines of the proof of Proposition 2 in [1]. (This part is actually a key observation for the solution to the cap set problem; it was discussed and largely formalized at the start of the formalization project.)

⋮

We continue after the definition of the $A \times A$ matrix B.

For all $m \in M_n^{d/2}$ the matrices $(m(a)F_m(b))_{(a,b)\in A\times A}$ and $(G_m(a)m(b))_{(a,b)\in A\times A}$ have rank ≤ 1 by Lemma 8.1. B is the sum of $2|M_n^{d/2}| = 2m_{d/2}$ of these rank ≤ 1 matrices, so by Proposition 7.3 (and induction) we get

$$\operatorname{rank}(B) \le 2m_{d/2}. \tag{7}$$

We note hat B is a diagonal matrix, indeed for $a, b \in A$ with $a \ne b$ we have $B_{a,b} = P(\alpha a + \beta b) = 0$ (first identity by definition of B and second by assumption 6). For $a \in A$ we have $B_{a,a} = P(\alpha a + \beta a) = P(-\gamma a)$ (first identity by definition of B and second by assumption 5). So by Lemma 8.2

$$\operatorname{rank}(B) = |\{\, a \in A \mid P(-\gamma a) \ne 0 \,\}|. \tag{8}$$

Together, (7) and (8) immediately imply the required conclusion.

12. Proof of Theorem 4 from [1]

Theorem 12.1 (Theorem 4 from [1])**.** *Let $\alpha, \beta, \gamma \in \mathbb{F}_q$ such that*

$$\alpha \neq 0 \text{ or } \beta \neq 0 \text{ or } \gamma \neq 0 \tag{9}$$

and

$$\alpha + \beta + \gamma = 0; \tag{10}$$

and let $A \subset \mathbb{F}_q^n$ such that

$$\forall a_1, a_2, a_3 \in A : \alpha a_1 + \beta a_2 + \gamma a_3 = 0 \Rightarrow a_1 = a_2 = a_3. \tag{11}$$

Then

$$|A| \leq 3m_{(q-1)n/3}.$$

Let α, β, γ, A satisfy (9), (10), and (11). By (9), at least one of α, β, γ is nonzero. Note that the whole theorem is symmetric in α, β, γ, so by swapping γ with α or β we can and will assume that

$$\gamma \neq 0. \tag{12}$$

Let

$$d \in [0, (q-1)n] \tag{13}$$

(in fact, we shall later only need $d = 2(q-1)n/3$) and let

$$V := \{\, p \in S_n^d \mid \forall a \in \mathbb{F}_q^n - (-\gamma A) : p(a) = 0 \,\}. \tag{14}$$

Lemma 12.2.

$$\dim(V) \geq m_d - q^n + |A|$$

Proof. Note that

$$|\mathbb{F}_q^n - (-\gamma A)| = |\mathbb{F}_q^n| - |-\gamma A| = q^n - |-\gamma A| = q^n - |A|,$$

where in the last step we used $|-\gamma A| = |A|$ since $A \to -\gamma A : a \mapsto -\gamma a$ is a bijection (since $\gamma \neq 0$). By Lemma 9.2 (with V, A, X replaced by $S_n^d, \mathbb{F}_q^n - (-\gamma A), V$) we have

$$\dim(V) \geq \dim(S_n^d) - |\mathbb{F}_q^n - (-\gamma A)| = m_d - (q^n - |A|) = m_d - q^n + |A|.$$

□

For any $p \in V$ we define the set

$$\mathrm{Sup}(p) := \{\, a \in \mathbb{F}_q^n \mid p(a) \neq 0 \,\}$$

Choose $P \in V$ such that

$$\forall P' \in V : \mathrm{Sup}(P) \subset \mathrm{Sup}(P') \Rightarrow \mathrm{Sup}(P) = \mathrm{Sup}(P'), \tag{15}$$

which is possible since $\mathbb{F}_q^n$ is finite (if such a P would not exist, then we could inductively obtain an arbitrarily long strictly increasing chain of subsets of $\mathbb{F}_q^n$, which is impossible). Let

$$\Sigma := \mathrm{Sup}(P).$$

Lemma 12.3.

$$|\Sigma| \geq \dim(V)$$

Proof. Suppose not, i.e. $|\Sigma| < \dim(V)$. Let

$$W := \{\, Q \in V \mid \forall a \in \Sigma : Q(a) = 0 \,\}$$

By Lemma 9.2

$$\dim(W) \geq \dim(V) - |\Sigma| > 0.$$

So we have some $Q \in W$ with $Q \neq 0$.

Now let $a \in \Sigma$, then by definition $Q(a) = 0$ and $P(a) \neq 0$, so

$$(P+Q)(a) = P(a) + Q(a) = P(a) \neq 0.$$

Since $Q \neq 0$, $\forall a \in \Sigma : Q(a) = 0$, and the associated function $e(Q) : \mathbb{F}_q^n \to \mathbb{F}_q$ is not identically zero, we have some $b \in \mathbb{F}_q^n - \Sigma$:

$$Q(b) \neq 0.$$

Since $\sup(P) = \Sigma$, we get

$$P(b) = 0,$$

hence

$$(P+Q)(b) = P(b) + Q(b) = Q(b) \neq 0.$$

We conclude that $P + Q \in V$ satisfies

$$\sup(P) = \Sigma \subsetneq \Sigma \cup \{b\} \subset \sup(P+Q),$$

which contradicts the choice of P (15). □

Lemma 12.4.

$$|\Sigma| \leq 2m_{d/2}$$

Proof. Let

$$S(A) := \{\, \alpha a_1 + \beta a_2 \in \mathbb{F}_q^n \mid a_1, a_2 \in A \text{ with } a_1 \neq a_2 \,\}.$$

We claim that

$$S(A) \cap -\gamma A = \emptyset. \tag{16}$$

Indeed, suppose the claim does not hold, then there are $a_1, a_2, a_3 \in A$ with $a_1 \neq a_2$ and

$$\alpha a_1 + \beta a_2 = -\gamma a_3.$$

From (11) we get $a_1 = a_2 = a_3$, contradicting $a_1 \neq a_2$, which proves the claim.

Note that any $p \in V$ vanishes on the complement of $-\gamma A$ by definition of V (14), so this holds in particular for $P \in V$. By (16) $S(A)$ is contained in the complement of $-\gamma A$, so P vanishes on $S(A)$, i.e. for all $a_1, a_2 \in A$ with $a_1 \neq a_2$:

$$P(\alpha a_1 + \beta a_2) = 0.$$

Applying Proposition 11.1 now gives

$$|\{\, a \in A \mid P(-\gamma a) \neq 0 \,\}| \leq 2m_{d/2}.$$

Together with $\Sigma = \{\, b \in \mathbb{F}_q^n \mid P(b) \neq 0 \,\} \subset -\gamma A$ we get (through the bijection $a \mapsto b := -\gamma a$)

$$|\Sigma| = |\{\, a \in A \mid P(-\gamma a) \neq 0 \,\}| \leq 2m_{d/2}.$$

□

Lemma 12.5.

$$q^n - m_d \leq m_{(q-1)n-d}$$

Proof. Let $I := \{\, x \in \mathbb{Z} \mid 0 \leq x \leq q-1 \,\}$ (so $|I| = q$). Note that we have a (canonical) bijection

$$f : M_n \to I^n, \quad \prod_{i=1}^{n} x_i^{a_1} \mapsto (a_1, \ldots, a_n).$$

Define

$$\begin{aligned}
A :=& I^n \\
B :=& \{\, (a_1, \ldots, a_n) \in I^n \mid \sum_{i=1}^{n} a_i \leq d \,\} \\
C :=& \{\, (a_1, \ldots, a_n) \in I^n \mid \sum_{i=1}^{n} a_i > d \,\} \\
D :=& \{\, (a_1, \ldots, a_n) \in I^n \mid \sum_{i=1}^{n} a_i < (q-1)n - d \,\} \\
E :=& \{\, (a_1, \ldots, a_n) \in I^n \mid \sum_{i=1}^{n} a_i \leq (q-1)n - d \,\}.
\end{aligned}$$

Note

$$\begin{aligned}
&|A| = q^n \\
&|B| = |f(M_n^d)| = |M_n^d| = m_d \\
&|C| = |A| - |B| = q^n - m_d \quad \text{(since } A = B \cup C \text{ and } B \cap C = \emptyset\text{)}.
\end{aligned}$$

Note that we have a bijection (actually an involution)

$$\iota : A \to A : \quad (a_1, \ldots, a_n) \to (q-1-a_1, \ldots, q-1-a_n).$$

And $\iota(C) = D$, since for $(a_1, \ldots, a_n) \in A$:

$$\sum_{i=1}^{n} a_i > d \Leftrightarrow \sum_{i=1}^{n} ((q-1) - a_i) = (q-1)n - \sum_{i=1}^{n} a_i < (q-1)n - d.$$

Note $D \subset E$, hence $|D| \leq |E|$. Putting things together gives

$$q^n - m_d = |C| = |\iota(C)| = |D| \leq |E| = |f(M_n^{(q-1)n-d})| = |M_n^{(q-1)n-d}| = m_{(q-1)n-d}.$$

□

Combining the four previous lemmata yields a nice upper bound for $|A|$.

Lemma 12.6.

$$|A| \leq 2m_{d/2} + m_{(q-1)n-d}$$

Proof.

$$\begin{aligned}
m_d - q^n + |A| &\leq \dim(V) && \text{(by Lemma 12.2)} \\
&\leq |\Sigma| && \text{(by Lemma 12.3)} \\
&\leq 2m_{d/2} && \text{(by Lemma 12.4)},
\end{aligned}$$

hence

$$m_d - q^n + |A| \leq 2m_{d/2}.$$

First adding $(q^n - m_d)$ to both sides, and next applying Lemma 12.5 gives

$$|A| \leq 2m_{d/2} + (q^n - m_d) \leq 2m_{d/2} + m_{(q-1)n-d}.$$

□

To finish the proof of Theorem 12.1, we take

$$d := 2(q-1)n/3,$$

which obviously meets (13). Now

$$d/2 = (q-1)n/3 \text{ and also } (q-1)n - d = (q-1)n/3,$$

so Lemma 12.6 specializes to

$$|A| \leq 3m_{(q-1)n/3},$$

which was to be proven.

13. Proof of asymptotics for $m_{(q-1)n/3}$

This includes Corollary 5 in[1].

For the asymptotics, [1] uses Cramér's theorem on large deviations. A more elementary approach, using Stirlings's approximation for the factorial function, was described in [2]. In [3] another (more) elementary approach, using recurrence sequences, was given. Inspired by *loc. cit.*, we work out yet another (even more elementary) approach, as follows.

We start with the geometric sum.

Proposition 13.1. *Let R be a ring, $x \in R$, and $m \in \mathbb{Z}_{\geq 0}$. Then*

$$(x-1)\sum_{j=0}^{m-1} x^j = x^m - 1.$$

Proof. Straightforward induction. (Note that for the case $m = 0$, the sum is empty, which has value 0.) □

Corollary 13.2. *Let $m \in \mathbb{Z}_{\geq 0}$ and $x \in k$. Then*

$$\sum_{j=0}^{m-1} x^j = \begin{cases} \frac{x^m-1}{x-1} & \text{if } x \neq 1; \\ m & \text{if } x = 1. \end{cases}$$

A basic estimate for the reciprocal geometric sum.

Lemma 13.3. *Let $x \in \mathbb{R}$ with $0 < x < 1$ and $M \in \mathbb{Z}_{\geq 0}$. Then*

$$\sum_{j=0}^{M} \frac{1}{x^j} \leq \frac{1}{1-x}\frac{1}{x^M}.$$

Proof.

$$\begin{aligned}\sum_{j=0}^{M}\frac{1}{x^j} &= \frac{1}{x^M}\sum_{j=0}^{M}\frac{x^M}{x^j}\\ &= \frac{1}{x^M}\sum_{j=0}^{M}x^{M-j}\\ &= \frac{1}{x^M}\sum_{i=0}^{M}x^{i}\\ &= \frac{1}{x^M}\frac{1-x^{M+1}}{1-x}\\ &< \frac{1}{x^M}\frac{1}{1-x}.\end{aligned}$$

□

Definition 13.4. Let $m \in \mathbb{Z}_{\geq 1}$. An element of k is called a *primitive m-th root of unity* if it has order m in the multiplicative group k^*.

Lemma 13.5. *Let $m \in \mathbb{Z}_{\geq 1}$, $\zeta \in k$ a primitive m-th root of unity, and $i \in \mathbb{Z}$. Then*

$$\sum_{j=0}^{m-1}\zeta^{ij} = \begin{cases}0 & \text{if } m \nmid i;\\ m & \text{if } m \mid i.\end{cases}$$

Proof. Apply Corollary 13.2 with $x = \zeta^i$, using that $(\zeta^i)^m = (\zeta^m)^i = 1^i = 1$ and that (by Lemma 4.1) $\zeta^i = 1$ if and only if $m \mid i$. □

This lemma is the basis for the discrete analog of using Cauchy's integral theorem for picking out a coefficient of a (Laurent) polynomial. We start with the constant coefficient.

Proposition 13.6. *Let $f \in k[x, x^{-1}]$ be a Laurent polynomial. Then for any positive integer $m > \max(\deg_x(f), \deg_{x^{-1}}(f))$, $\zeta \in k$ a primitive m-th root of unity, and $r \in k^*$, the constant coefficient of mf is given by*

$$\sum_{j=0}^{m-1} f(r\zeta^j).$$

Proof. Write $f = \sum_{i=-(m-1)}^{m-1} a_i x^i$ for (unique) $a_{-(m-1)}, \ldots, a_{m-1} \in k$ (any of which could equal zero). So the constant coefficient of mf is ma_0. Then

$$\begin{aligned}\sum_{j=0}^{m-1} f(r\zeta^j) &= \sum_{j=0}^{m-1}\sum_{i=-(m-1)}^{m-1} a_i(r\zeta^j)^i && \text{(by substituting definitions)}\\ &= \sum_{i=-(m-1)}^{m-1} a_i r^i \sum_{j=0}^{m-1}\zeta^{ij} && \text{(by changing order of summation, etc.)}\\ &= a_0 r^0 m && \text{(by Lemma 13.5, using that for } i = -(m-1), \ldots, m-1 :\ m \mid i \Leftrightarrow i = 0)\\ &= ma_0 && \text{(by commutativity and since } r^0 = 1 \text{ by definition)}\end{aligned}$$

□

As a quick corollary we have a handy formula for picking out any coefficient. The proof of the previous proposition could of course also be adapted to prove the theorem below directly (so that we do not have to introduce Laurent polynomials).

Theorem 13.7. *Let $f \in k[x]$ and $i \in \mathbb{Z}_{\geq 0}$. Then for any integer $m > \max(\deg(f), i)$, $\zeta \in k$ a primitive m-th root of unity, and $r \in k^*$, the i-th coefficient of mf is given by*

$$\sum_{j=0}^{m-1} \frac{f(r\zeta^j)}{r^i \zeta^{ij}}.$$

Proof. Note that f/x^i is a Laurent polynomial whose constant coefficient equals the i-th coefficient of f. Now apply Proposition 13.6 with f/x^i instead of f. □

Define for $j \in \mathbb{Z}_{\geq 0}$

$$c_j^{(n)} := |\{ (a_1, \ldots, a_n) \in \{0, 1, \ldots, q-1\}^n \mid \sum_{i=1}^{n} a_i = j \}|.$$

We have an obvious expression for m_d as sum of these values.

Lemma 13.8.

$$m_d = \sum_{j=0}^{\lfloor d \rfloor} c_j^{(n)}.$$

Proof. This follows immediately by noting, as in the proof of Lemma 12.5, that

$$m_d = |\{ (a_1, \ldots, a_n) \in \{0, 1, \ldots, q-1\}^n \mid \sum_{i=1}^{n} a_i \leq d \}|.$$

□

In line with [3], we recognise the $c_j^{(n)}$'s as certain polynomial coefficients.

Lemma 13.9. *For all $j \in \mathbb{Z}_{\geq 0}$: $c_j^{(n)}$ equals the j-th coefficient of the polynomial*

$$\left(1 + x + \ldots + x^{q-1}\right)^n \in \mathbb{Z}[x].$$

Proof. For 'fixed ' q, the statement follows by induction on n. The Lemma trivially holds for $n = 1$. Now suppose the lemma holds for a certain $n \in \mathbb{Z}_{>0}$. Denote $f := 1 + x + \ldots + x^{q-1} \in \mathbb{Z}[x]$. Now for all $j \in \mathbb{Z}_{\geq 0}$:

$$\mathrm{coef}_j(f^{n+1}) = \mathrm{coef}_j(f^n f) = \sum_{i=0}^{j} \mathrm{coef}_i(f^n)\, \mathrm{coef}_{j-i}(f) = \sum_{i=0}^{j} c_i^{(n)} c_{j-i}^{(1)} = c_j^{(n+1)},$$

where the last equality follows readily (on paper...) from the definition of the $c_J^{(N)}$'s. □

Using our earlier theorem to pick out coefficient, we can easily estimate $c_j(n)$ (and hence m_d).

Lemma 13.10. *For any $j \in \mathbb{Z}_{\geq 0}$ and any $r \in \mathbb{R}_{>0}$:*

$$c_j^{(n)} \leq \frac{(1 + r + \ldots + r^{q-1})^n}{r^j}.$$

Proof. Choose $m \in \mathbb{Z}_{>0}$ with $m > \max(n(q-1), j)$. Let $\zeta := \exp(2\pi\sqrt{-1}/m) \in \mathbb{C}$. Then ζ is a primitive m-th root of unity and $|\zeta| = 1$. Now

$$\begin{aligned}
c_j^{(n)} &= \operatorname{coef}_j\left((1 + x + \ldots + x^{q-1})^n\right) && \text{(by Lemma 13.9)}\\
&= \frac{1}{m}\sum_{i=0}^{m-1} \frac{(1 + r\zeta^j + \ldots + (r\zeta^j)^{q-1})^n}{r^j\zeta^{ij}} && \text{(by Theorem 13.7)}\\
&= \frac{1}{m}\left|\sum_{i=0}^{m-1} \frac{(1 + r\zeta^j + \ldots + (r\zeta^j)^{q-1})^n}{r^j\zeta^{ij}}\right| && \text{(since } c_j^{(n)} \geq 0\text{)}\\
&\leq \frac{1}{m}\sum_{i=0}^{m-1} \frac{(|1| + |r\zeta^j| + \ldots + |r\zeta^j|^{q-1})^n}{|r^j\zeta^{ij}|} && \text{(by the triangle inequality, etc.)}\\
&= \frac{1}{m}\sum_{i=0}^{m-1} \frac{(1 + r + \ldots + r^{q-1})^n}{r^j} && \text{(since } |\zeta| = 1 \text{ and } r > 0\text{)}\\
&= \frac{(1 + r + \ldots + r^{q-1})^n}{r^j} && \text{(since all } m \text{ summands are independent of } i\text{).}
\end{aligned}$$

□

For $r \in \mathbb{R}_{>0}$ and $q \in \mathbb{Z}_{>0}$ write

$$C_{r,q} := \frac{1 + r + \ldots + r^{q-1}}{r^{(q-1)/3}}.$$

For simplicity we first estimate the relevant asymptotics for Theorem 12.1 in case $3|n$ (which is harmless.)

Lemma 13.11. *Let $n = 3N$ with $N \in \mathbb{Z}_{>0}$ and $r \in \mathbb{R}$ with $0 < r < 1$. Then*

$$m_{(q-1)n/3} = m_{(q-1)N} \leq \frac{1}{1-r} C_{r,q}^n.$$

Proof.

$$\begin{aligned}
m_{(q-1)N} &= \sum_{j=0}^{(q-1)N} c_j^{(3N)} && \text{(by Lemma 13.8)}\\
&\leq \sum_{j=0}^{(q-1)N} \frac{(1 + r + \ldots + r^{q-1})^{3N}}{r^j} && \text{(by Lemma 13.10)}\\
&= (1 + r + \ldots + r^{q-1})^{3N} \sum_{j=0}^{(q-1)N} \frac{1}{r^j}\\
&\leq (1 + r + \ldots + r^{q-1})^{3N} \frac{1}{r^{(q-1)N}} \frac{1}{1-r} && \text{(by Lemma 13.3)}\\
&= \frac{1}{1-r}\left(\frac{1 + r + \ldots + r^{q-1}}{r^{(q-1)/3}}\right)^n && \text{(since } N = n/3\text{)}\\
&= \frac{1}{1-r} C_{r,q}^n && \text{(by definition).}
\end{aligned}$$

□

We want to consider all $n \in \mathbb{Z}_{>0}$ but do not care about the constant $1/(1-r)$.

Remark 13.12. Below we *do not* use Landau's *big O* notation since all multiplicative constants can easily be made explicit.

Theorem 13.13. *Let $n \in \mathbb{Z}_{>0}$ and $r \in \mathbb{R}$ with $0 < r < 1$. Then*

$$m_{(q-1)n/3} \leq \frac{C_{r,q}^2}{1-r} C_{r,q}^n.$$

Proof. Let $i \in \{0,1,2\}$ such that $n' := n + i$ is divisible by 3. Then

$$\begin{aligned}
m_{(q-1)n/3} &\leq m_{(q-1)n'/3} && \text{(since } d_1 \leq d_2 \Rightarrow m_{d_1} \leq m_{d_2}\text{; similarly for increasing } n) \\
&\leq \frac{1}{1-r} C_{r,q}^{n'} && \text{(by Lemma 13.11 with } n' \text{ instead of } n) \\
&\leq \frac{1}{1-r} C_{r,q}^{n+2} && \text{(since } n' \leq n+2 \text{ and } C_{r,q} > 1) \\
&= \frac{C_{r,q}^2}{1-r} C_{r,q}^n && \text{(by rewrite).}
\end{aligned}$$

□

Together with Theorem 12.1 we arrive at the following explicit asymptotic upper bound for $|A|$.

Theorem 13.14. *Let the notation and conditions be as in Theorem 12.1 and let $r \in \mathbb{R}$ with $0 < r < 1$. Then*

$$|A| \leq \frac{3C_{r,q}^2}{1-r} C_{r,q}^n.$$

Proof.

$$\begin{aligned}
|A| &= 3m_{(q-1)n/3} && \text{(by Theorem 12.1)} \\
&\leq \frac{3C_{r,q}^2}{1-r} C_{r,q}^n && \text{(by Theorem 13.13).}
\end{aligned}$$

□

Lemma 13.15. *For all $q \in \mathbb{Z}_{\geq 2}$ there exists an $r \in \mathbb{R}$ with $0 < r < 1$ such that $C_{r,q} < q$.*

Proof. Note $C_{1,q} = q$ and $\frac{d}{dr} C_{r,q}|_{r=1} > 0 \ldots$
OR show via elementary way that with $r := (q-1)/(q+2)$ we have $C_{r,q} < q \ldots$ □

We get our main result.

Theorem 13.16. *Let the notation and conditions be as in Theorem 12.1. There exist $B, C \in \mathbb{R}_{>0}$ with $C < q$, such that for all $n \in \mathbb{Z}_{>0}$: $|A| \leq B \cdot C^n$.*

Proof. Choose r as in Lemma 13.15 such that $C := C_{r,q} < q$, and let $B := 3C_{r,q}^2/(1-r)$. Now apply Theorem 13.14. □

For several values of q we can be very explicit about a value of c that works. (In fact, the values of c below will be optimal (or a precise numerical approximation thereof) given the proof method, but we shall not prove this.) We keep the notation and conditions as in Theorem 12.1.

First the *cap set problem*.

Theorem 13.17. *Let $q = 3$ and $c := \frac{3}{8} \sqrt[3]{207 + 33\sqrt{33}}$. Then $|A| = \mathcal{O}(c^n)$.*

Proof. Let $r := (-1+\sqrt{33})/8$. Then $0 < r < 1$ and

$$C_{r,q} = \frac{1+r+r^2}{r^{2/3}} = \frac{3}{8}\sqrt[3]{207+33\sqrt{33}}.$$

Now apply Theorem 13.14. □

From now on we express results numerically for convenience. For $3 \le q < 50$ a prime power we define real (or even rational) numbers r_q and C_q in the following table. (Note that the case $q = 2$ is not very interesting.)

q	r_q	C_q
3	0.59307033081725358248132643352736616477753305724785	2.75510461302363300022127654653368619521587542646 35
4	0.65729810613837599082505552000480171164504768618926	3.61071861327603934981864900838405862746510858573 65
5	0.70415268204813468689198669322132696343638075791281	4.46157776570257781140848213884764090297903871883 05
7	0.76780677969242521445924402966336695842991975939993	6.15620486321673841642864499054684146620540582787 76
8	0.79038872370790461105610875566178974381236959429548	7.00155475499400745813224731239145313714129088115 78
9	0.80897722790523491111266988703268198209152068465472	7.84612058258580571250658432756124593073923758112 29
11	0.83776747968601146197311784198457216654957997594351	9.53368539207555099275584600143742953300044116813 0
13	0.85902510166765698759178640475123322423462991338793	11.21990798911487743710354307874670920930982759504
16	0.88218894242913276503686171517072556839234842435080	13.74776213458745700186503241050408124151866114255
17	0.88830758043507257326234095105382783282727016343919	14.59011716296566909733292591731068890488059646602
19	0.89881856557451588672455240338174322453086526354899	16.27455106840026451533064402457870391069290751032
23	0.91484764548356435179270332283816978457151143428432	19.64263645872880846855649800712949925601009365863
25	0.92109805634831861028657723785713430394845587760601	21.32640831010467458504116449092661287095567773974
27	0.92649385906623403043220219721359816616215863083796	23.01005118248578718050749133276735336886865749163
29	0.93119907005099645802948397596620351689310355515557	24.69359086763659075194245991317129618483238935200
31	0.93533826774922682502465725097793299160275771076073	26.37704670973149142375844507925916831824336087591
32	0.93722658990392746349985428712960304612407039759198	27.21874795251220672602238447846984018710029446169
37	0.94522485380783942198322445203198854044829554690688	31.42704353767073321497532153760864868680225675 64
41	0.95029187505308080110749234851246282509920205764457	34.79348741912559389651285702611249403969615890 56
43	0.95248941722836481804826949752866830065974434752949	36.47666120987899740665519079203246794153653113 33
47	0.95634899796299730225165301786027369278180052980541	39.84293254975881186354589016121037533477372541 86
49	0.95805283290333388408037131751427604295954501076278	41.52603619692476777873321962510612007396867876 36

Theorem 13.18. *Let $3 \le q < 50$ be a prime power. Then $|A| = \mathcal{O}\left(C_q^n\right)$.*

Proof. By Theorem 13.14 we have $|A| = \mathcal{O}\left(C_{r_q,q}^n\right)$. Now check that $C_{r_q,q} \le C_q$. *Not yet checked for round off errors*, perhaps the last digit of some of the C_q's has to be raised by 1 ... □

These values agree with those given in [3] (where $4 \le q \le 31$).

Remark 13.19. Exact (optimal, for the method) values for C_4 and C_5 could be given is terms of radicals, analogous to Theorem 13.17.

14. A Further simplification for the asymptotics

Thanks to a remark from Dion Gijswijt on our preprint we can significantly simplify the asymptotics proof from the previous section.

Recall the following two definitions and two lemmas from the previous section.

For $j \in \mathbb{Z}_{\ge 0}$ let

$$c_j^{(n)} := |\{\,(a_1,\ldots,a_n) \in \{0,1,\ldots,q-1\}^n \mid \sum_{i=1}^n a_i = j\,\}|.$$

For $r \in \mathbb{R}_{>0}$ and $q \in \mathbb{Z}_{>0}$ write

$$C_{r,q} := \frac{1+r+\ldots+r^{q-1}}{r^{(q-1)/3}}.$$

Lemma (Lemma 13.8)**.**

$$m_d = \sum_{j=0}^{\lfloor d \rfloor} c_j^{(n)}.$$

Lemma (Lemma 13.9)**.** *For all $j \in \mathbb{Z}_{\geq 0}$: $c_j^{(n)}$ equals the j-th coefficient of the polynomial*

$$\left(1 + x + \ldots + x^{q-1}\right)^n \in \mathbb{Z}[x].$$

We quickly obtain a slightly improved version of Theorem 13.13 (and Lemma 13.11).

Theorem 14.1. *Let $n \in \mathbb{Z}_{>0}$ and $r \in \mathbb{R}$ with $0 < r < 1$. Then*

$$m_{(q-1)n/3} \leq C_{r,q}^n.$$

Proof. Write $e := \lfloor (q-1)n/3 \rfloor$. Note that for integers $0 \leq j \leq e$ (using $0 < r < 1$)

$$r^e \leq r^j. \tag{17}$$

Now

$$\begin{aligned}
m_{(q-1)n/3} r^e &= \sum_{j=0}^{e} c_j^{(n)} r^e && \text{(by Lemma 13.8 and distributivity)} \\
&\leq \sum_{j=0}^{e} c_j^{(n)} r^j && \text{(using (17) and that the } c_j^{(n)} \geq 0) \\
&\leq \sum_{j=0}^{(q-1)n} c_j^{(n)} r^j && \text{(since } e < (q-1)n \text{ and all summands are} \geq 0) \\
&= \left(1 + r + \ldots + r^{q-1}\right)^n && \text{(by Lemma 13.9).}
\end{aligned}$$

Hence

$$m_{(q-1)n/3} \leq \frac{\left(1 + r + \ldots + r^{q-1}\right)^n}{r^{\lfloor (q-1)n/3 \rfloor}} \leq \frac{\left(1 + r + \ldots + r^{q-1}\right)^n}{r^{(q-1)n/3}} = \left(\frac{1 + r + \ldots + r^{q-1}}{r^{(q-1)/3}}\right)^n = C_{r,q}^n.$$

□

Using Theorem 14.1 instead of Theorem 13.13 we can replace Theorems 13.14 and 13.16 by the following slightly cleaner versions.

Theorem 14.2. *Let the notation and conditions be as in Theorem 12.1 and let $r \in \mathbb{R}$ with $0 < r < 1$. Then*

$$|A| \leq 3C_{r,q}^n.$$

Theorem 14.3. *Let the notation and conditions be as in Theorem 12.1. There exists $C \in \mathbb{R}_{>0}$ with $C < q$, such that for all $n \in \mathbb{Z}_{>0}$: $|A| \leq 3 \cdot C^n$.*

References

[1] Jordan S. Ellenberg and Dion Gijswijt, On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression, *Ann. of Math.* (2) 185 (2017), no. 1, 339–343.

[2] Terence Tao, A symmetric formulation of the Croot-Lev-Pach-Ellenberg-Gijswijt capset bound, *Blog*, May 18, 2016, `https://terrytao.wordpress.com/2016/05/18/a-symmetric-formulation-of-the-croot-lev-pach-ellenberg-gijswijt-capset-bound/`.

[3] Doron Zeilberger, A Motivated Rendition of the Ellenberg-Gijswijt Gorgeous proof that the Largest Subset of F_3^n with No Three-Term Arithmetic Progression is $O(c^n)$, with $c = \sqrt[3]{(5589 + 891\sqrt{33})/8} = 2.755104613023633000221 27\ldots$, *preprint*, July 6, 2016, `https://arxiv.org/abs/1607.01804`.