The MDS Conjecture

Joseph Briggs

December 5, 2014

Simeon Ball: "On subsets of $\mathbb{F}_q^k$ of which every set of size k is a basis" (2010)

Throughout, we will assume $q = p^h$ for some prime p.

Definition. Let $S \subset \mathbb{F}_q^k$. Then S is *generic* if no k such vectors of S lie in a hyperplane, ie all subsets of size k form a basis. (In the projectivised space, S is often called an *arc*.)

The *MDS conjecture* states that every generic set $S \subset \mathbb{F}_k^q$ has size $|S| \leq q + 1$, unless $(p = 2 \wedge k \in \{3, q-1\})$ or $k > q$.
Henceforth, we assume $S \subset \mathbb{F}_p^k$ is generic.

Definition. The reason for the title MDS is because of the tie with *maximum distance separable* codes. Indeed, a $[n, k, d]$ linear code C is said to be *maximum distance separable* if it meets the Singleton Bound $k \leq n - d + 1$ at equality. This bound is simply proved by noting that in a code of distance d, any choice of $n - d + 1$ entries can determine at most one codeword of C, so in the case of equality, any choice must yield some codeword.

Lemma. *Given an MDS $[n, k, d]$ code $U \leq \mathbb{F}_q^n$ we can construct a generic $S \subset \mathbb{F}_q^k$ of size n, and vice versa.*

Proof. Given S, let $U \leq \mathbb{F}_q^{|S|}$ be the space generated by the rows $u_1, \ldots, u_k$ of the matrix A whose columns are S. (A has full rank $\Rightarrow \dim U = k$.) Then if we have some nonzero element of U

$$u = \sum_{\delta \in \Delta} c_\delta u_\delta,$$

then u having k entries $= 0$ is equivalent to saying that the corresponding k elements of S are in the hyperplane $\{\sum_{\delta\in\Delta} c_\delta X_\delta = 0\}$, which contradicts the genericity of S. Thus U as a code has minimum weight (and hence d) $\geq |S| - k + 1$, so is an MDS code! Conversely, given U, write a basis for U as the rows of some matrix A, and take S to be the n columns. Then we can reverse the above argument: If k of the columns were in the hyperplane $\{\sum_{\delta\in\Delta} c_\delta X_\delta = 0\}$, these columns would give k entries of $\sum_{\delta\in\Delta} c_\delta u_\delta$ which were 0. So $d \leq n - k$, contradicting U being an MDS code. So in fact any k elements of S are LI. □

Note. Using the above result, sizes of generic sets are equivalent to lengths of MDS codes. The most famous class of codes meeting the Singleton Bound are Reed Solomon codes, where we encode the space of all polynomials of degree $\leq q-1$ via $Enc: f \mapsto (f(a_1), \ldots, f(a_n))$. This encoder needs to use *distinct* a_i, so we can obtain a length of up to $n = q$ by using all possible elements of $\mathbb{F}_q$.
By using the basis $\{1, t, t^2, \ldots, t^{k-1}\}$ for the rows of A in the above proof, and we obtain the following generic set.

Example 1. We can attain the bound of $q+1$ by choosing

$$S = \{(1, t, t^2, \ldots, t^{k-1}) : t \in \mathbb{F}_q\} \cup \{(0, 0, \ldots, 0, 1) = e_k\}$$

(The first k vectors are called the *normal rational curve* of $\mathbb{F}_q^k$. We've actually been able to add e_k to the vectors merely obtained from the RS code, and this can be viewed colloquially as allowing $t = \infty$ on the curve as well. This concept can be made rigorous when working with projective space.)
Then $|S| = |\mathbb{F}_q| + 1 = q + 1$. Though we already knew that any k vectors from the curve are LI by construction, we can check explicitly here that they form a $k \times k$ Vandermonde matrix, whose determinant is therefore $\neq 0$. Similarly, any collection of $k-1$ vectors together with e_k form the matrix $\begin{pmatrix} & V & & \mathbf{0} \\ t_1^{k-1} & t_2^{k-1} & \ldots & t_{k-1}^{k-1} & 1 \end{pmatrix}$ where V is a $(k-1)\times(k-1)$ VDM, hence has det $\neq 0$. So S is generic, as desired.
What is curious about this example is how the dual of RS codes being RS codes translates to a similar self-dual property of this S. More specifically, write $n = q + 1$. If we list S as the columns of a matrix G and pick a full rank $(n-k)\times n$ matrix H such that $HG^T = 0$, then the n columns of H form a generic set S' (we will see this later). In fact, for this specific S, we

have

$$G = \begin{pmatrix} 1 & 1 & \cdots & 1 & 0 \\ t_1 & t_2 & \cdots & t_q & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ t_1^{k-2} & t_2^{k-2} & \cdots & t_q^{k-2} & 0 \\ t_1^{k-1} & t_2^{k-1} & \cdots & t_q^{k-1} & 1 \end{pmatrix}, H = \begin{pmatrix} 1 & 1 & \cdots & 1 & 0 \\ t_1 & t_2 & \cdots & t_q & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ t_1^{q-k-1} & t_2^{q-k-1} & \cdots & t_q^{q-k-1} & 0 \\ t_1^{q-k} & t_2^{q-k} & \cdots & t_q^{q-k} & 1 \end{pmatrix}$$

is (one) valid choice for H, since it is full rank for the same reason G is, and any inner product of two of their rows

$$\begin{aligned}(HG^T)_{ij} =& < H_i, G_j > = \sum_{t \in \mathbb{F}_q} t^i t^j \\ &= \sum_{m=0}^{q-2} g^{m(i+j)} = \frac{g^{(q-1)(i+j)} - 1}{g^{i+j} - 1} = 0\end{aligned}$$

where $g \in \mathbb{F}_q$ is primitive, and we assume $i + j < q - 1$ (ie we haven't chosen the bottom row twice-in this case, note $(\sum t^{q-1}) + 1 = (q-1) + 1 = 0$ in $\mathbb{F}_q$ too). Then taking S' to be the columns of H gives us back the same example!

Example 2. (Glynn,1986)
Let

$$S = \{(1, x, x^2 + \eta x^6, x^3, x^4) : x \in \mathbb{F}_9\} \cup \{e_5\}$$

Here, $\eta^4 = -1$ is a constant (and, in $\mathbb{F}_9$, there are 4 such choices of η). Then certainly $|S| = q + 1 = 10$, and using multilinearity of det:

$$0 = \det \begin{pmatrix} 1 & \dots & 1 \\ t_1 & \dots & t_5 \\ t_1^2 + \eta t_1^6 & \dots & t_5^2 + \eta t_5^6 \\ t_1^3 & \dots & t_5^3 \\ t_1^4 & \dots & t_5^4 \end{pmatrix} = \eta \det \begin{pmatrix} \mathbf{1} \\ \mathbf{t} \\ \mathbf{t^6} \\ \mathbf{t^3} \\ \mathbf{t^4} \end{pmatrix} + \det \begin{pmatrix} \mathbf{1} \\ \mathbf{t} \\ \mathbf{t^2} \\ \mathbf{t^3} \\ \mathbf{t^4} \end{pmatrix}$$

$$\Rightarrow -\eta^3 \det \begin{pmatrix} \mathbf{1} \\ \mathbf{t^3} \\ \mathbf{t^2} \\ \mathbf{t} \\ \mathbf{t^4} \end{pmatrix} = \det \begin{pmatrix} \mathbf{1} \\ \mathbf{t^3} \\ \mathbf{t^6} \\ \mathbf{t} \\ \mathbf{t^4} \end{pmatrix} = -\det \begin{pmatrix} \mathbf{1} \\ \mathbf{t} \\ \mathbf{t^6} \\ \mathbf{t^3} \\ \mathbf{t^4} \end{pmatrix} = \eta^{-1} \det \begin{pmatrix} \mathbf{1} \\ \mathbf{t} \\ \mathbf{t^2} \\ \mathbf{t^3} \\ \mathbf{t^4} \end{pmatrix}$$

.

Here $\mathbf{t^i}$ denotes $(t_1^i, t_2^i, t_3^i, t_4^i, t_5^i)$. (Note that the implication results from cubing the previous line, where we note that characteristic $3 \Rightarrow (a+b)^3 = a^3 + b^3 \forall a, b \in \mathbb{F}_9$ and det is a polynomial $\Rightarrow \det(X)^3 = det(X^3)$ for X as a list of k^2 variables.) So, if the t_i are distinct, this is a VDM matrix with $\neq 0$ determinant, and so $\eta^3 = \eta^{-1} \Rightarrow \eta^4 = 1$. Contradiction! So the original det was $\neq 0$, and we have linear independence.
Choosing e_5 in place of $(1, t_5, t_5^2 + \eta t_5^6, t_5^3, t_5^4)$ results in a similar proof of linear independence, noting that the $\mathbf{t^4}$'s will be removed from the above determinants.

Example 3. For the first exception to the conjecture (ie $(p = 2 \wedge k \in \{3, q-1\})$), we note that S can actually have size $q+2$, by adding $e_{k-1} = (0, \ldots, 0, 1, 0)$ to S from Example 0.2. Here, any $k-$subset using both e_{k-1} *and* e_k yields a $(k-2) \times (k-2)$ VDM matrix, so has $\neq 0$ determinant. so we need only check that $k-1$ vectors on the normal rational curve, together with e_{k-1}, are LI.
For $k = 3$, note that

$$\det \begin{pmatrix} 1 & 1 & 0 \\ \alpha & \beta & 1 \\ \alpha^2 & \beta^2 & 0 \end{pmatrix} = \alpha^2 - \beta^2 = (\alpha - \beta)^2 \neq 0$$

(crucially over characteristic 2) since α and β are distinct. Hence the 3 vectors are LI.
For $k = q-1$, note that

$$\det \begin{pmatrix} 1 & \cdots & 1 & 0 \\ \alpha_1 & \cdots & \alpha_{k-1} & 0 \\ \vdots & & \vdots & \vdots \\ \alpha_1^{k-2} & \cdots & \alpha_{k-1}^{k-2} & 1 \\ \alpha_1^{k-1} & \cdots & \alpha_{k-1}^{k-1} & 0 \end{pmatrix} = \pm \left(\sum_{i=1}^{q-2} \alpha_i \right) \prod_{i<j} (\alpha_i - \alpha_j)$$

Here, the formula for the determinant comes from noticing all $(\alpha_i - \alpha_j)$ must appear as linear factors (as with the VDM derivation), but the resulting polynomial is homogeneous of degree $\binom{q-2}{2}$. By contrast, the det is homogeneous of degree $1 + 2 + \cdots + (q-3) + (q-1) = \binom{q-2}{2} + 1$, so there is an additional linear factor. By symmetry, it must be a nonzero scaling of $\sum \alpha_i$.
Thus, it suffices to check that this term is also $\neq 0$, but note that (using

$q > 2$):

$$0 = [X^{q-1}](X^q - X) = [X^{q-1}] \prod_{\alpha \in \mathbb{F}_q} (X - \alpha) = -\sum_{\alpha \in \mathbb{F}_q} \alpha$$

$$\Rightarrow \sum_{i=1}^{q-2} \alpha_i = \beta + \gamma \neq 0$$

where β, γ are the only 2 elements of $\mathbb{F}_q$ not appearing as α_i for any i (and are certainly distinct).

Example 4. (Bush,1952) For the second exception, we consider $k > q$, and note that $|S| = k + 1 > q + 1$ is attainable via

$$S = \{(\lambda_1, 0, \ldots, 0), (0, \lambda_2, \ldots, 0), \ldots, (0, 0, \ldots, \lambda_k), (1, 1, \ldots, 1)\}$$

for any collection of $\lambda_i \neq 0$. All that is required from the $(k+1)-$th vector is that all entries are $\neq 0$, since any hyperplane containing $k - 1$ of the first vectors is of the form $X_i = 0$ (if the $i-$th was excluded).

Stability of Maximal Examples

We can view 2 generic sets as *equivalent* if they can be mapped to one another via automorphisms of $\mathbb{F}_q$ and/or changes of basis in $\mathbb{F}_q^k$.

Proposition. *For $k \geq q$, any generic S with $|S| \geq k+1$ in fact has $|S| = k+1$ and is equivalent to example 4.*

Proof. First take any $k+1$ vectors of S, and note they have a linear dependence $\sum_{i=1}^{k+1} c_i S^{(i)} = 0$. None of the c_i's are 0, else we would have the other k $S^{(i)}$'s lying in a hyperplane. So we can choose our basis to be $e_i = -\frac{c_i}{c_{k+1}} S^{(i)}$ for $1 \leq i \leq k$, noting the first $k S^{(i)}$ are LI. Then $S^{(1)} = (-\frac{c_{k+1}}{c_1}, 0, \ldots, 0), \ldots, S^{(k)} = (0, \ldots, 0, -\frac{c_{k+1}}{c_k})$ wrt basis $(e_i)_{i=1}^k$ and $S^{(k+1)} = (1, 1, \ldots, 1)$ by the given dependence.
Now, any additional vector v in S has all nonzero entries (by the same logic as for $S^{(k+1)}$). So all k of them are in $\mathbb{F}_q^\times$. By pigeonhole, 2 of them are the same, wlog $v_1 = v_2$. But now $\{X_1 - X_2 = 0\}$ is a hyperplane containing the k vectors $S^{(3)}, \ldots, S^{(k)}, S^{(k+1)}, v$, contradicting genericity-so there was no such additional v. □

Note. Stability for the extremal examples in the MDS conjecture is not true in general-Example 2 was the first different construction to Example 1 for q odd and can be shown to be not equivalent using techniques from projective geometry. There is another example due to Hirschfeld (1971) where $(0, 0, 0, 1)$ is added to the curve $\{(1, t, t^{2^r}, t^{2^r+1})\}$ for any $(r, h) = 1$ where $q = 2^h$, which has also been proven distinct. Also, for the parameter choices as in Example 3 allowing for a size $q+2$ generic set, many examples have been constructed, all of which are not equivalent to the hyperoval. However, some stability is known (and indeed proved by Ball in the same paper), in particular for the small case $k \leq p$ (as per the proof below) and the large case $q - p + 1 \leq k \leq q - 2$. Naturally, none of the above "counterexamples" satisfy these conditions!

It should be noted, however, that all known examples are some extension of Reed-Solomon Codes (under the aforementioned correspondence with MDS codes), though not all necessarily encoding the collection of polynomials of degree $\leq k - 1$ (as illustrated by e.g. Example 2).

Proof of the MDS conjecture for $k \leq p$ (hence for $p = q$)

Definition. (Segre, 1967) Let $Z \subset S, |Z| = k-2$. Consider the codimension-1 hyperplanes $\Sigma \supset Z$ defined by $\Sigma = \{x : f_\Sigma(x) = 0\}$ for linear functions f (effectively, for normal vectors v_Σ, have $f_\Sigma(x) =< x, v_\Sigma >$). Then we define the *tangent function* on $\mathbb{F}_q^k$ to be the multivariate polynomial

$$T_Z(X) := \prod_{\Sigma : \Sigma \cap S = Z} f_\Sigma(X)$$

Note. We observe that, if $H =< Z >$, then $\dim H = k-2$ (otherwise adding any 2 vectors gives a dim $\leq k-1$ space). So $|H| = q^{k-2}$, and since the Σ (1 dimension higher) containing V only intersect in V (they are uniquely determined by any additional vector), they each have $q^{k-1} - |V|$ of the $q^k - |V|$ remaining vectors, and hence there are precisely $q+1$ such Σ. Each such Σ has at most 1 additional element of S (other than those from Z), since S is generic. Let t be the number of such Σ with *no* extra element of S, so there are t (linear) terms in the product $T_Z(X)$. Then because $|S \backslash Z|$ is the number of such Σ with 1 element of S, we actually know $t + |S \backslash Z| = q+1$ is the total such number of hyperplanes Σ, and hence we know $t = q + k - 1 - |S|$. (Crudely $t \geq 0 \Rightarrow |S| \leq k+q-1$ is a starting weak bound, but we are aiming for $t \geq k-2$.)

Lemma. *(Segre) Suppose that $Y \cup \{x, y, z\} \subset S$ is a collection of k distinct elements (so certainly LI by genericity of S). Then*

$$T_{Y \cup \{x\}}(y) T_{Y \cup \{y\}}(z) T_{Y \cup \{z\}}(x) = (-1)^{t+1} T_{Y \cup \{x\}}(z) T_{Y \cup \{y\}}(x) T_{Y \cup \{z\}}(y)$$

Proof. Consider $\{x, y, z\} \cup Y = \{e_1, e_2, \ldots, e_k\} = B$ as a basis for $\mathbb{F}_q^k$. Let $Z = Y \cup z$. Consider the $q-1$ hyperplanes $\{\alpha X_1 - X_2 = 0\} : \alpha \in \mathbb{F}_q^\times\}$ together with $\{X_1 = 0\}$ and $\{X_2 = 0\}$. They are all distinct, and all contain $Z = \{e_3, \ldots, e_k\}$. In fact, the latter two contain y and x respectively. Certainly, they make up all such $q+1$ hyperplanes $\Sigma \supset Z$.
Let $D = S \backslash B$. Then any $d \in D$ has $d_1 \neq 0 \neq d_2$ (else it would be in the planes containing y or x), and $< Z, d >$ is the plane $\{d_2 X_1 - d_1 X_2 = 0\}$.
Knowing the general form of all hyperplanes containing Z and being able to describe those containing another element of S, we can now get a handle on those without (and hence T_Z). Specifically, for any f_Σ in the T_Z product,

$\Sigma = \{X : f_\Sigma(X) = v_{\Sigma 1}X_1 + v_{\Sigma 2}X_2 = 0\}$. So t of the $\alpha \in \mathbb{F}_q^\times$ are of the form $\frac{-v_{\Sigma 1}}{v_{\Sigma 2}}$, for $\Sigma \cap S = Z$, and the remaining $q - 1 - t = |S| - k$ such are of the form $\frac{d_2}{d_1}$, for $d \in D$.
Notice that $x_1 = 1 = y_2, x_2 = 0 = y_1$, so

$$\begin{aligned}\frac{T_Z(x)}{T_Z(y)} &= \frac{\prod_{\Sigma \cap S = Z}(1.v_{\Sigma 1} + 0)}{\prod_{\Sigma \cap S = Z}(0 + 1.v_{\Sigma 2})} \\ &= (-1)^t \prod_{d \in D} \frac{d_1}{d_2} \prod_{\Sigma \cap S = Z} \frac{-v_{\Sigma 1}}{v_{\Sigma 2}} \prod_{d \in D} \frac{d_2}{d_1} \\ &= (-1)^t \prod_{d \in D} \frac{d_1}{d_2} \prod_{\alpha \in \mathbb{F}_q^\times} \alpha = (-1)^{t+1} \prod_{d \in D} \frac{d_1}{d_2}.\end{aligned}$$

By applying similar logic with $Z = Y \cup \{y\}$ and $Z = Y \cup \{z\}$, deduce

$$\begin{aligned}\frac{T_{Y\cup\{x\}}(y)T_{Y\cup\{y\}}(z)T_{Y\cup\{z\}}(x)}{T_{Y\cup\{x\}}(z)T_{Y\cup\{y\}}(x)T_{Y\cup\{z\}}(y)} &= (-1)^{3t+3} \prod_{d \in D} \frac{d_2}{d_3}\frac{d_3}{d_1}\frac{d_1}{d_2} \\ &= (-1)^{t+1}.\end{aligned}$$

□

Note. It should be noted that this was a new proof of Segre's original lemma. The original proof relied heavily on working with all coordinates X_i at once, whereas here, we have saved a lot of work with the initial choice of basis.

We can obtain another relation involving the tangent function by viewing it as a polynomial and using interpolation:

Lemma. *For $E = \{a_1, \ldots, a_{t+2}\}$ and $|Y| = k - 2$ disjoint in S,*

$$0 = \sum_{a \in E} T_Y(a) \prod_{b \in E \setminus a} \det(a, b, Y)^{-1}.$$

Note that since $|Y| = k - 2$, the determinant is of a $k \times k$ matrix, and the vectors are distinct elements of S and hence LI by genericity.

Proof. Firstly, for any $x \in \mathbb{F}_q^k$,

$$T_Y(x) = \sum_{i=1}^{t+1} T_Y(a_i) \prod_{l \neq i, t+2} \frac{\det(x, a_l, Y)}{\det(a_i, a_l, Y)}$$

Indeed, setting $x = a_j$ for some $1 \leq i \leq t+1$, we see whenever $i \neq j$ that

$$\prod_{l \neq i, t+2} \det(a_j, a_l, Y) = 0$$

since the term $l = j$ has a determinant with a_j repeated. So the only term of the sum surviving is for $a_i = a_j$, and the term is precisely $T_Y(a_j)$ because of the scaling. So the two polynomials agree on a set of $t+1$ points, and since they are of degree t, must be equal.
Now let $x = a_{t+2}$. Dividing through by

$$\prod_{l=1}^{t+1} \det(a_{t+2}, a_l, Y)$$

kills the numerator of each product in the sum, and also adds $\det(a_{t+2}, a_i, Y) = -\det(a_i, a_{t+2}, Y)$ to the denominator of the $i-$th term of the sum. Thus

$$T_Y(a_{t+2}) \prod_{l=1}^{t+1} \det(a_{t+2}, a_l, Y)^{-1} = -\sum_{i=1}^{t+1} T_Y(a_i) \prod_{l \neq i} \det(a_i, a_l, Y)^{-1},$$

as desired (with $b = a_l, a = a_i$). □

Note. Studying the proof carefully above, it seems as though we've only used a very weak version of the genericity of $E \cup Y \subset S$-we always pick $k-$sets which fully contain Y, and only use 2 elements of E at a time-the argument above does not even detect that no 3 elements of E are dependent.
So, a major idea of this proof is to swap elements of Y and E one-by-one, so the resulting interpolation equation encapsulates more of the strength of the genericity of S.

Theorem. *Say Y and E are as before. Write $Y = \{y_1, \ldots, y_{k-2}\}.$, and $\theta_i = (a_1, \ldots, a_{i-1}, y_i, \ldots, y_{k-2})$, as a set as well as a tuple. Then, for any "valid" r,*

$$0 = \sum_{\{a_1, \ldots, a_r\} \in \binom{E}{r}} \left(\prod_{i=1}^{r-1} \frac{T_{\theta_i}(a_i)}{T_{\theta_{i+1}}(y_i)} \right) T_{\theta_r}(a_r) \prod_{z \in E \cup Y \setminus (\theta_r \cup \{a_r\})} \frac{1}{\det(a_r, z, \theta_r)}.$$

In addition, all $r!$ terms in the sum corresponding to reorderings of $a_1, \ldots, a_r \in E$ have the same value, so we have in fact

$$0 = r! \sum_{a_1 < \cdots < a_r \in E} \left(\prod_{i=1}^{r-1} \frac{T_{\theta_i}(a_i)}{T_{\theta_{i+1}}(y_i)} \right) T_{\theta_r}(a_r) \prod_{z \notin \theta_r \cup \{a_r\}} \frac{1}{\det(a_r, z, \theta_r)}.$$

Note. For r to be "valid" for the above to make sense, we need $r \leq k-1$, as $\theta_{k-1} = (y_1, \ldots, y_{k-2})$ is the largest such well-defined θ_r. (We can take $r > t+2 = |E|$ but then the above sum is empty and the result is vacuous.) This is a genuine generalisation of the interpolation equation, since taking $r = 1$ makes the first product empty, $T_{\theta_1}(a_1) = T_Y(a_1)$ noting $\theta_1 = Y$ as sets, and z now plays the role of a_l in the second product.

Proof. (half-omitted) We show the equality of terms corresponding to rearrangements of a_i's.
In fact, we need only show it for swapping a_j and $a_{j+1}\, \forall j$, since the transpositions $(j\, j+1)$ generate the symmetric group S_{t+2}. Now, let $(a'_i)_{i=1}^{t+2}$ be the sequence given by swapping a_j and a_{j+1}, with corresponding θ'_i's. Then $\theta_i = \theta'_i$ as sets for every i except $i = j$, where they have a_j and a_{j+1} respectively. For now, assume $j < r-1$. Then

$$\prod_{i=1}^{r-1} \frac{T_{\theta_i}(a_i) T_{\theta'_{i+1}}(y_i)}{T_{\theta_{i+1}}(y_i) T_{\theta'_i}(a'_i)} = \frac{T_{\theta'_j}(y_{j-1}) T_{\theta_j}(a_j) T_{\theta_{j+1}}(a_{j+1})}{T_{\theta_j}(y_{j-1}) T_{\theta'_j}(a'_j) T_{\theta'_{j+1}}(a'_{j+1})}$$
$$= \frac{T_{\theta'_j}(y_{j-1}) T_{\theta_j}(a_j) T_{\theta_{j+1}}(a_{j+1})}{T_{\theta_j}(y_{j-1}) T_{\theta'_j}(a_{j+1}) T_{\theta_{j+1}}(a_j)} = (-1)^{t+1}$$

by Segre's Lemma, using $\{x, y, z\} = \{a_j, a_{j+1}, y_{j-1}\}$. As a tuple, θ'_r is θ_r with a_j and a_{j+1} swapped, so also

$$\prod_{z \notin \theta_r \cup \{a_r\}} \frac{\det(a_r, z, \theta_r)}{\det(a_r, z, \theta'_r)} = \prod_{z \in E \cup Y \setminus (\theta_r \cup \{a_r\})} -1 = (-1)^{(k+t-(k-2)-1)} = (-1)^{t+1}.$$

So the ratio of the term from sequence (a_i) to the term from the transposed sequence (a'_i) is $\frac{(-1)^{2t+2} T_{\theta_r}(a_r)}{T_{\theta'_r}(a'_r)} = 1$ using $\theta'_r = \theta_r$ as sets and $a_r = a'_r$.
The same argument works for $j = r-1$ with $a'_{r-1} = a_r, a'_r = a_{r-1}$, noting

$$\frac{T_{\theta_r}(a_r)}{T_{\theta'_r}(a'_r)} \prod_{i=1}^{r-1} \frac{T_{\theta_i}(a_i) T_{\theta'_{i+1}}(y_i)}{T_{\theta_{i+1}}(y_i) T_{\theta'_i}(a'_i)} = (-1)^{t+1}$$

by Segre's Lemma with $\{x, y, z\} = \{a_{r-1}, a_r, y_{r-1}\}$, and

$$\det(a'_r, z, \theta'_r) = \det(a_{r-1}, z, a_1, \ldots, a_{r-2}, a_r, y_r, \ldots)$$
$$= -\det(a_r, z, a_1, \ldots, a_{r-2}, a_{r-1}, y_r, \ldots) = -\det(a_r, z, \theta_r).$$

To prove the equation one merely needs to use induction on r: indeed, we already have case $r = 1$. There are no additional ideas required, but using the equality of terms corresponding to the same r-sets now proven can simplify the notation a great deal. Roughly speaking, one has to apply the induction hypothesis once for each $b \in E$ that can be used as the new a_i (swapping out for y_r). □

Theorem. *Assuming $(t+2)+(k-2) \leq |S|$ and (for the first time) $k \leq p$, we have $t \geq k-2$-the desired improvement on crude bound $q+k-1-|S| = t \geq 0$ previously mentioned-and hence $|S| \leq q+1$.*

Proof. Say for contradiction $t \leq k-3$. By the bounds on S, we may choose disjoint E and Y of sizes $t+2, k-2$; and $r = t+2 \leq k-1$ is valid for applying the previous theorem.

$$0 = (t+2)! \left(\prod_{i=1}^{t+1} \frac{T_{\theta_i}(a_i)}{T_{\theta_{i+1}}(y_i)} \right) T_{\theta_r}(a_r) \prod_{z \notin \theta_{t+1} \cup \{a_{t+1}\}} \frac{1}{\det(a_{t+1}, z, \theta_{t+1})}$$

All terms in this product except for $(t+2)!$ are$\in \mathbb{F}_q^\times$, so in fact $(t+2)! = 0 \Rightarrow p|(t+2)! \Rightarrow t \geq p-2 \geq k-2$. Contradiction! □

Theorem. *Suppose $t + k > |S|$, and as before, $k \leq p$. Then $|S| \leq q+1$.*

Proof. We apply a "duality transformation" $S \to S'$, where $|S'| = |S|$ is $n-k = k'$-generic, as follows: listing the n elements of S, let

$$G = \left(\begin{array}{ccccc} S^{(1)} & | & \ldots & | & S^{(n)} \end{array} \right) \} k$$

and choose a full rank $(n-k) \times n$ matrix H such that $HG^T = 0$.
Then $\ker H = G^T(\mathbb{F}_q^n)$: $\supset$ is clear and both have dim $= k$. Then we may take S' to be the n columns of H. This S' is generic because any $(n-k)$-linear dependence in the columns of H says precisely $Hx = 0$ for some $x \neq 0$ of weight $\leq n-k$, but $x = G^T y$ has weight $\geq n-k+1$ otherwise $\geq k$ of the vectors $S^{(i)}$ lie in the hyperplane $< X, y >= 0$, contradicting genericity of S.
Now, let $t' = q + k' - 1 - |S'| = q - 1 - k$. Symmetrically $t = q - 1 - k'$.
Either $|S'| \geq t' + k'$ and we can apply the previous result again to obtain $|S| = |S'| \leq q+1$, or $|S| = |S'| < t' + k'$ and $|S| < t + k \Rightarrow |S| < \frac{t'+k+t+k'}{2} = q - 1$. □

Note. The last 2 results combine to give the full proof of the MDS conjecture in the case $k \leq p$. More generally for $k < q$, the same argument shows $|S| \leq q + k - 1 - \min\{k, p\}$. A more recent paper of Ball (2011) actually extends this proof to include $k \leq 2p - 2$ when $q > p$, and this is currently the best known.