Permutation Groups and the Strength of Quantum Finite Automata with Mixed States *

Rūsiņš Freivalds, Māris Ozols, Laura Mančinska

Institute of Mathematics and Computer Science, University of Latvia,
Raiņa bulvāris 29, Rīga, Latvia

Abstract. It was proved earlier by A. Ambainis and R. Freivalds that the quantum finite automata with pure states can have exponentially smaller number of states than the deterministic finite automata recognizing the same language. There is a never published "folk theorem" claiming that the quantum finite automata with mixed states are no more than super-exponentially concise than the deterministic finite automata. It is not known whether the super-exponential advantage of the quantum automata with mixed states is really achievable.
We show how this problem can be reduced to a certain problem about permutation groups, namely: (1) if there is a fixed constant c and an infinite sequence of distinct integers n such that for each n there is a group G_n of permutations of the set $\{1, 2, \ldots, n\}$ such that $|G_n| = e^{\Omega(n \log n)}$ and the pairwise Hamming distance of permutations is at least $c \cdot n$, then (2) there is an infinite sequence of languages L_n such that for each language there is a quantum finite automata with mixed states that recognizes the language L_n and has $O(n)$ states, while any deterministic finite automaton recognizing L_n must have at least $e^{\Omega(n \log n)}$ states.
We do not know whether (1) is true, but we provide a list of several known results on permutation groups, that possibly could be used to prove (1). However, note that if (1) turns out to be false, it does not imply, that (2) is also false.

1 Introduction

A. Ambainis and R. Freivalds proved in [1] that for recognition of some languages the quantum finite automata can have smaller number of states than the deterministic ones, and this difference can even be exponential. The proof contained a slight non-constructiveness, and the exponent was not shown explicitly. For probabilistic finite automata exponentiality of such a distinction was not yet proved. The best (smaller) gap was proved by Ambainis [2]. The languages recognized by automata in [1] were presented explicitly but the exponent was not. In a very recent paper by R.Freivalds [3] the non-constructiveness is modified, and an explicit (and seemingly much better) exponent is obtained at the expense of having only non-constructive description of the languages used. Moreover, the best estimate in this paper was proved under the assumption of the well-known

* This research is supported by Grant No.05.1528 from the Latvian Council of Science.

Artin's Conjecture (1927) in Number Theory. [3] contains also a theorem that does not depend on any open conjectures but the estimate is worse, and the description of the languages used is even less constructive. This seems to be the first result in finite automata depending on open conjectures in Number Theory.

The following two theorems are proved in [3]:

Theorem 1. *Assume Artin's Conjecture. There exists an infinite sequence of regular languages $L_1, L_2, L_3, \ldots$ in a 2-letter alphabet and an infinite sequence of positive integers $z_1, z_2, z_3, \ldots$ such that for arbitrary j:*

(1) there is a probabilistic reversible automaton with z_j states that recognizes the language L_j with the probability $\frac{19}{36}$,
(2) any deterministic finite automaton recognizing L_j has at least $(2^{\frac{1}{4}})^{z_j} = (1.189207115\ldots)^{z_j}$ states.

Theorem 2. *There exists an infinite sequence of regular languages $L_1, L_2, L_3, \ldots$ in a 2-letter alphabet and an infinite sequence of positive integers $z_1, z_2, z_3, \ldots$ such that for arbitrary j:*

(1) there is a probabilistic reversible automaton with z_j states that recognizes the language L_j with the probability $\frac{68}{135}$,
(2) any deterministic finite automaton recognizing L_j has at least $(7^{\frac{1}{14}})^{z_j} = (1.149116725\ldots)^{z_j}$ states.

The two theorems above are formulated in [3] as assertions about reversible probabilistic automata. For probabilistic automata (reversible or not) it was unknown before the paper [3] whether the gap between the size of probabilistic and deterministic automata can be exponential. It is easy to re-write the proofs in order to prove counterparts of Theorems 1 and 2 for quantum finite automata with pure states. The aim of this paper is to propose a way how to prove a counterpart of these theorems for quantum finite automata with mixed states.

2 Quantum automata with mixed states

Quantum algorithms with mixed states were first considered by D. Aharonov, A. Kitaev, N. Nisan [4]. More detailed description of quantum finite automata with mixed states can be found in A. Ambainis, M. Beaudry, M. Golovkins, A. Ķikusts, M. Mercer, D. Thérien [5]. Since we are interested only in the most simple and the most restricted version of these automata, we consider only so-called Latvian QFA in this paper. These are the quantum finite automata that can be implemented using the Nuclear Magnetic Resonance (NMR) technology. All the other types of quantum finite automata with mixed states are less restrictive.

The automaton is defined by the initial density matrix ρ_0. Every symbol a_i in the input alphabet is associated with a unitary matrix U_i. When the automaton

reads the symbol s_i, the current density matrix ρ is transformed into $U_i \rho U_i^\dagger$. When the reading of the input word is finished and the end-marker "$" is read, the current density matrix ρ is transformed into $U_\$ \rho U_\$^\dagger$ and separate measurements of all states are performed. After that the probabilities of all the accepting states are totaled, and the probabilities of all the rejecting states are totaled.

It is easy to see that quantum finite automata with pure states described by C. Moore and J. Crutchfield [23] but not the quantum finite automata with pure states described by A. Kondacs and J. Watrous [22] can be simulated by Latvian QFA with no increase in the number of states.

3 From quantum automata to permutations

In this section we show how the problem of proving that quantum finite automata with mixed states have a super-exponential advantage over deterministic automata can be reduced to a certain problem about permutation groups.

Definition 1. *The* Hamming distance *or simply* distance $d(r,s)$ *between two* n*-permutations* r *and* s *on the set* S *is the number of elements* $x \in \mathrm{S}$ *such that* $r(x) \neq s(x)$*. The* similarity $e(r,s)$ *is the number of* $x \in \mathrm{S}$ *such that* $r(x) = s(x)$*. Note that* $d(r,s) + e(r,s) = |\mathrm{S}| = n$*.*

Theorem 3. *The assertion (1) implies the assertion (2), where:*

(1) there is a fixed constant c *and an infinite sequence of distinct integers* n *such that for each* n *there is a group* G_n *of permutations of the set* $\{1, 2, \ldots, n\}$*, the group has* $e^{\Omega(n \log n)}$ *elements and* k *generating elements, and the pairwise Hamming distance of permutations is at least* $c \cdot n$*,*

(2) there is an infinite sequence of distinct integers n *such that for each* n *there is a language* L_n *in a* k*-letter alphabet that can be recognized with probability* $\frac{c}{2}$ *by a quantum finite automata with mixed states that has* $2n$ *states, while any deterministic finite automaton recognizing* L_n *must have at least* $e^{\Omega(n \log n)}$ *states.*

Proof. For each permutation group G_n we define the language L_n as follows:

The letters of L_n *are the* k *generators of the group* G_n *and it consists of words* $s_1 s_2 s_3 \ldots s_m$ *such that the product* $s_1 \circ s_2 \circ s_3 \circ \cdots \circ s_m$ *differs from the identity permutation.*

We will construct a quantum automaton with mixed states. It has $2n$ states and the initial density matrix ρ_0 is a diagonal block-matrix that consists of n blocks $\widetilde{\rho}_0$:

$$\widetilde{\rho}_0 = \frac{1}{2n} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. \tag{1}$$

For example, in the case $n = 4$ the density matrix ρ_0 is given in (3).

For each of k generators $g_i \in G_n$ we will construct the corresponding unitary matrix U_i as follows – it is a $2n \times 2n$ permutation matrix, that permutes the

elements in the even positions according to permutation g_i, but leaves the odd positions unpermuted.

For example, $g = 3241$ can be expressed as the following permutation matrix that acts on a column vector:

$$g = \begin{pmatrix} 0\,0\,1\,0 \\ 0\,1\,0\,0 \\ 0\,0\,0\,1 \\ 1\,0\,0\,0 \end{pmatrix}. \tag{2}$$

The initial density matrix ρ_0 for $n = 4$ and the unitary matrix U that corresponds to the permutation matrix (2) of premutation g are as follows:

$$\rho_0 = \frac{1}{8} \begin{pmatrix} 1\,1\,0\,0\,0\,0\,0\,0 \\ 1\,1\,0\,0\,0\,0\,0\,0 \\ 0\,0\,1\,1\,0\,0\,0\,0 \\ 0\,0\,1\,1\,0\,0\,0\,0 \\ 0\,0\,0\,0\,1\,1\,0\,0 \\ 0\,0\,0\,0\,1\,1\,0\,0 \\ 0\,0\,0\,0\,0\,0\,1\,1 \\ 0\,0\,0\,0\,0\,0\,1\,1 \end{pmatrix}, \quad U = \begin{pmatrix} 1\,0\,0\,0\,0\,0\,0\,0 \\ 0\,0\,0\,0\,0\,1\,0\,0 \\ 0\,0\,1\,0\,0\,0\,0\,0 \\ 0\,0\,0\,1\,0\,0\,0\,0 \\ 0\,0\,0\,0\,1\,0\,0\,0 \\ 0\,0\,0\,0\,0\,0\,0\,1 \\ 0\,0\,0\,0\,0\,0\,1\,0 \\ 0\,1\,0\,0\,0\,0\,0\,0 \end{pmatrix}. \tag{3}$$

The unitary matrix $U_\$$ for the end-marker is also a diagonal block-matrix. It consists of n blocks that are the *Hadamard matrices*

$$\widetilde{H} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{4}$$

Notice how the Hadamard matrix $\widetilde{H}$ acts on two specific 2×2 density matrices:

$$\text{if } \rho = \frac{1}{2n} \begin{pmatrix} 1\,1 \\ 1\,1 \end{pmatrix}, \text{ then } \widetilde{H}\rho\widetilde{H}^\dagger = \frac{1}{2n} \begin{pmatrix} 2\,0 \\ 0\,0 \end{pmatrix}, \tag{5}$$

$$\text{if } \rho = \frac{1}{2n} \begin{pmatrix} 1\,0 \\ 0\,1 \end{pmatrix}, \text{ then } \widetilde{H}\rho\widetilde{H}^\dagger = \frac{1}{2n} \begin{pmatrix} 1\,0 \\ 0\,1 \end{pmatrix}. \tag{6}$$

For example, when the letter g is read, the unitary matrix U is applied to the density matrix ρ_0 (both are given in equation (3)) and the density matrix $\rho_1 = U\rho_0 U^\dagger$ is obtained. When the end-marker "\$" is read, the density matrix becomes $\rho_\$ = U_\$ \rho_1 U_\$^\dagger$. Matrices ρ_1 and $\rho_\$$ are as follows:

$$\rho_1 = \frac{1}{8} \begin{pmatrix} 1\,0\,0\,0\,0\,0\,0\,1 \\ 0\,1\,0\,0\,1\,0\,0\,0 \\ 0\,0\,1\,1\,0\,0\,0\,0 \\ 0\,0\,1\,1\,0\,0\,0\,0 \\ 0\,1\,0\,0\,1\,0\,0\,0 \\ 0\,0\,0\,0\,0\,1\,1\,0 \\ 0\,0\,0\,0\,0\,1\,1\,0 \\ 1\,0\,0\,0\,0\,0\,0\,1 \end{pmatrix}, \quad \rho_\$ = \frac{1}{8} \begin{pmatrix} 1 & 0 & 0\,0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & 0\,0 & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 2\,0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0\,0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0\,0 & 1 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & 0\,0 & 0 & 1 & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0\,0 & \frac{1}{2} & -\frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 0\,0 & \frac{1}{2} & -\frac{1}{2} & 0 & 1 \end{pmatrix}. \tag{7}$$

Finally, we declare the states in the even positions to be accepting, but the states in the odd positions to be rejecting. Therefore one must sum up the diagonal entries that are in the even positions of the final density matrix to find the probability that a given word is accepted.

In our example the final density matrix $\rho_\$$ is given in (7). It corresponds to the input word "$g\$$", which is accepted with probability $\frac{1}{8}(1+0+1+1) = \frac{3}{8}$ and rejected with probability $\frac{1}{8}(1+2+1+1) = \frac{5}{8}$. Note that the accepting and rejecting probabilities sum up to 1.

It is easy to see, that the words that do not belong to the language L_n are rejected with certainty, because the matrix $U_\$\rho_0 U_\$^\dagger$ has all zeros in the even positions on the main diagonal. However, the words that belong to L_n are accepted with the probability at least $\frac{d}{2n} = \frac{cn}{2n} = \frac{c}{2}$, because all permutations are at least at the distance d from the identity permutation.

It is also easy to see that any deterministic automaton that recognizes the language L_n must have at least $N = |G_n|$ states, where $|G_n|$ is the size of the permutation group G_n. If the number of states is less than N, then there are two distinct words u and v such that the deterministic automaton ends up in the same state no matter which one of the two words it reads. Since G_n is a group, for each word we can find an inverse, that returns the automaton in the initial state (the only rejecting state). Since u and v are different, they have different inverses and $u \circ u^{-1}$ is the identity permutation and must be rejected, but $v \circ u^{-1}$ is not the identity permutation and must be accepted – a contradiction. □

4 Sharply transitive permutation groups

We are interested in permutation groups such that distinct permutations have large Hamming distance (see Sect. 3 for the definition of the *Hamming distance* $d(r,s)$ and the *similarity* $e(r,s)$ of two permutations r and s). It turns out that the notion of Hamming distance is related to the *multiple transitivity* of groups.

Definition 2. *A group G of permutations on the set* S *is called* k-transitive *if for every two k-tuples $(x_1, x_2, \ldots, x_k)$ and $(y_1, y_2, \ldots, y_k)$ of distinct elements of* S*, there is a permutation $p \in G$ such that $p(x_i) = y_i$ for all $i \in \{1, 2, \ldots, k\}$. If there is exactly one such permutation p, then G is called* sharply k-transitive*. Note that a sharply k-transitive group is also sharply $(k-1)$-transitive.*

It seems that it has been noted only recently (see, e.g., [10]) that the sharp k-transitivity imposes a restriction on the Hamming distance. This is given by the following trivial lemma:

Lemma 1. *If G is a sharply k-transitive set of n-permutations, then for any distinct $r, s \in G$:*

$$d(r,s) \geq n - k + 1. \tag{8}$$

Proof. Let us assume that $d(r,s) < n - k + 1$ for some $r, s \in G$. It means, the similarity $e(r,s) \geq k$ or both permutations act on some t-tuple ($t \geq k$) in the same way. This is a contradiction, since G is sharply k-transitive. □

Definition 3. *$G(n,d)$ denotes the size of the largest group of n-permutations with the pairwise distance at least d ($d \le n$).*

An analogue of the *Singleton bound* can be obtained for permutations [11]:

Lemma 2. *The following upper bound holds:*

$$G(n,d) \le \underbrace{n(n-1)(n-2)\cdots(d+1)d}_{n-d+1 \text{ multipliers}} \tag{9}$$

with equality if and only if there is a sharply $(n-d+1)$-transitive group of permutations.

Proof. We have $d(r,s) \ge d$ and $e(r,s) \le n-d$ for every distinct $r,s \in G$. It means, if we fix any $(n-d+1)$-tuple x and apply all permutations from G to it, the obtained tuples y must be different. The number of such tuples y can not exceed the right hand side of (9).

If the size of the group G matches the upper bound, then all possible tuples y of $n-d+1$ distinct elements of S can be obtained if all permutations of G are applied to any fixed tuple x. Moreover, for each y there is no more than one such permutation. It means we can send any $(n-d+1)$-tuple x to any y with exactly one permutation thus the group is sharply $(n-d+1)$-transitive.

The other way round – if the group is sharply $(n-d+1)$-transitive, we can send any fixed $(n-d+1)$-tuple x to any y with exactly one permutation. Thus there are at least as many permutations in the group as the right hand side of (9). There are no other permutations, otherwise it would be possible to send the given x to some y with two different permutations, which is a contradiction with the assumption that the group is sharply $(n-d+1)$-transitive. □

In the next several subsections we list the main known results on sharply k-transitive groups (see [6] for the basic facts, [7, 9] for additional information).

4.1 Sharply n-transitive and $(n-1)$-transitive groups

It is clear that the *symmetric group* S_n of all permutations on the set $\{1,2,\ldots,n\}$ is sharply n-transitive, because there is exactly one permutation that sends any n-tuple to any other n-tuple. However, S_n is also sharply $(n-1)$-transitive, because if the action of the permutation on $n-1$ elements is known, the action on the last element is uniquely determined. From Lemma 1 we obtain that the distance between distinct permutations of S_n is at least 2. It is clear, because distance 1 is not possible for permutations.

The group S_n can be generated by two generators (in cycle notation):

$$g_1 = (12)(3)(4)\ldots(n), \tag{10}$$

$$g_2 = (123\ldots n). \tag{11}$$

The first one corresponds to a transposition of first two elements, but the second one – to a cyclic shift of all elements. The group S_n consists of $n!$ permutations.

4.2 Sharply $(n-2)$-transitive groups

The *signature* or *sign* of a permutation s is defined as the parity of the number of *inversions* in s, i.e., pairs i, j such that $i < j$, but $s(i) > s(j)$. For example, $s = 3241$ has 4 inversions, namely 32, 31, 21, and 41 thus it is an even permutation. It is easy to show that a *transposition* (a permutation that swaps to elements) changes the sign of a permutation to the opposite. In fact the signature "sgn" of a permutation is a group homomorphism from S_n to $\{-1, 1\}$, because for any two permutations r and s we have $\operatorname{sgn}(s \circ r) = \operatorname{sgn} s \cdot \operatorname{sgn} r$. Therefore it is not hard to see that the set of all even permutations of the set $\{1, 2, \ldots, n\}$ forms a group – the *alternating group* A_n.

It is simple to show that A_n is sharply $(n-2)$-transitive – if we know the action of a permutation on $n-2$ elements, the remaining two elements can be either swapped or remain in the same order. One of these cases corresponds to an even permutation, but the other – to odd. From Lemma 1 it follows that the pairwise distance of distinct permutations of A_n is at least 3. This can also be obtained directly – even permutations can not have distance 2, because then they differ only by one transposition and therefore have different signs. Since the number of odd and even permutations is the same, A_n consists of $n!/2$ permutations.

4.3 Sharply 1-transitive groups

An example of a sharply 1-transitive group is the cyclic group C_n that consists of n permutations and is generated by a cyclic shift

$$g_1 = (123 \ldots n). \tag{12}$$

C_n is clearly sharply 1-transitive, because there is exactly one way how to shift any element to any other. The pairwise distance between distinct elements of C_n is exactly n.

4.4 Sharply 2-transitive groups

An infinite sequence of sharply 2-transitive groups can be constructed using an *affine transformation* $y(x) = ax + b$, where $a, b \in \mathbb{F}_n$, $a \neq 0$. Here $\mathbb{F}_n$ denotes the *finite field* of order n, i.e., a set of n elements together with two binary operations – addition and multiplication, such that $(\mathbb{F}_n, +)$ and $(\mathbb{F}_n \backslash \{0\}, *)$ are Abelian groups and both distribute laws hold. Such a field $\mathbb{F}_n$ exists if and only if n is a power of a prime number.

The function $y(x)$ acts on the elements of the field $\mathbb{F}_n$ as a permutation, because $ax_1 + b = ax_2 + b$ implies $x_1 = x_2$. There are in total $n(n-1)$ such permutations and they form a group: if $y_1(x) = a_1 x + b_1$ and $y_2(x) = a_2 x + b_2$, then $(y_1 \circ y_2)(x) = y_1(y_2(x)) = (a_1 a_2)x + (a_1 b_2 + b_1)$ which is also an affine transformation. To prove that the group is sharply 2-transitive, we have to show

that there is a unique solution to the following system of two linear equations:

$$\begin{cases} y_1 = ax_1 + b \\ y_2 = ax_2 + b \end{cases} \tag{13}$$

where $x_1 \neq x_2$ and $y_1 \neq y_2$. This solution is

$$a = \frac{y_1 - y_2}{x_1 - x_2}, \quad b = \frac{x_1 y_2 - x_2 y_1}{x_1 - x_2}. \tag{14}$$

Thus the group is sharply 2-transitive. In a similar way one can explicitly show that the pairwise Hamming distance between distinct permutations is at least $n - 1$, but it follows from Lemma 1 as well.

4.5 Sharply 3-transitive groups

Let $\mathbb{F}_q$ be a finite field, where q is a power of a prime number.

Definition 4. *The* multiplicative group $\mathbb{F}_q^*$ *of the field* $\mathbb{F}_q$ *is the set of all non-zero elements of* $\mathbb{F}_q$*, i.e.,* $F_q^* = \mathbb{F}_q \backslash \{0\}$.

Definition 5. *The* general linear group $\mathrm{GL}(m, \mathbb{F}_q)$ *is the set of all invertible* $m \times m$ *matrices with elements from the field* $\mathbb{F}_q$*. A matrix is* invertible *if it has a non-zero determinant.*

Definition 6. *The* projective general linear group $\mathrm{PGL}(m, \mathbb{F}_q)$ *is almost the same as* $\mathrm{GL}(m, \mathbb{F}_q)$*, except that matrices* $M, M' \in \mathrm{GL}(m, \mathbb{F}_q)$ *are treated as equal if there is a non-zero scalar* $c \in \mathbb{F}_q^*$ *such that* $M = cM'$*. In other words* $\mathrm{PGL}(m, q) = \mathrm{GL}(m, q)/\mathbb{F}_q^*$.

The matrices from the set $\mathrm{GL}(m, \mathbb{F}_q)$ form a group, because the matrix multiplication is associative, the product of two invertible matrices is also invertible and for each invertible matrix one can find an inverse. This group acts on the set of non-zero m-dimensional column vectors over $\mathbb{F}_q$ as a permutation. The set $\mathrm{PGL}(m, \mathbb{F}_q)$ consists of equivalence classes of matrices and is also a group. It acts on the equivalence classes of non-zero column vectors as a permutation.

The group $\mathrm{PGL}(m, \mathbb{F}_q)$ is sharply 3-transitive, when $m = 2$. To prove this, it is sufficient to show that for any two 3-tuples of vectors $X = (\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3)$ and $Y = (\mathbf{y}^1, \mathbf{y}^2, \mathbf{y}^3)$ from different equivalence classes there is exactly one matrix $M \in \mathrm{PGL}(2, \mathbb{F}_q)$ that sends the tuple X to Y. For vectors $\mathbf{x}^i$ and $\mathbf{y}^i$ this means

$$M \cdot \mathbf{x}^i = c_i \cdot \mathbf{y}^i, \tag{15}$$

where the constant c_i is introduced, because we are dealing with the equivalence classes of vectors. Thus for each $i \in \{1, 2, 3\}$ we have a linear system of equations

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1^i \\ x_2^i \end{pmatrix} = c_i \begin{pmatrix} y_1^i \\ y_2^i \end{pmatrix}. \tag{16}$$

We have to solve these systems with respect to matrix M, but in addition we have three unknown constants: c_1, c_2, and c_3. In fact, we are allowed to choose one of them and thus specify some particular matrix M from its equivalence class. If $c_3 = 1$, the three systems of linear equations (16) are equivalent to

$$\begin{pmatrix} x_1^1 & x_2^1 & 0 & 0 & y_1^1 & 0 \\ 0 & 0 & x_1^1 & x_2^1 & y_2^1 & 0 \\ x_1^2 & x_2^2 & 0 & 0 & 0 & y_1^2 \\ 0 & 0 & x_1^2 & x_2^2 & 0 & y_2^2 \\ x_1^3 & x_2^3 & 0 & 0 & 0 & 0 \\ 0 & 0 & x_1^3 & x_2^3 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ -c_1 \\ -c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ y_1^3 \\ y_2^3 \end{pmatrix}. \tag{17}$$

It has a unique solution if the determinant does not vanish. Using some algebraic manipulations one can show that

$$\begin{vmatrix} x_1^1 & x_2^1 & 0 & 0 & y_1^1 & 0 \\ 0 & 0 & x_1^1 & x_2^1 & y_2^1 & 0 \\ x_1^2 & x_2^2 & 0 & 0 & 0 & y_1^2 \\ 0 & 0 & x_1^2 & x_2^2 & 0 & y_2^2 \\ x_1^3 & x_2^3 & 0 & 0 & 0 & 0 \\ 0 & 0 & x_1^3 & x_2^3 & 0 & 0 \end{vmatrix} = \begin{vmatrix} x_1^1 & x_1^3 \\ x_2^1 & x_2^3 \end{vmatrix} \cdot \begin{vmatrix} x_1^2 & x_1^3 \\ x_2^2 & x_2^3 \end{vmatrix} \cdot \begin{vmatrix} y_1^1 & y_1^2 \\ y_2^1 & y_2^2 \end{vmatrix} \neq 0. \tag{18}$$

None of the three small determinants vanish, because we were given that the 3-tuples X and Y consist of vectors from different equivalence classes, but such vectors are clearly linearly independent. Therefore the big determinant clearly does not vanish as well and the system (17) has a unique solution.

Let us find the number of equivalence classes of vectors and matrices in $\mathrm{PGL}(2, \mathbb{F}_q)$. The number of 2-dimensional non-zero vectors over $\mathbb{F}_q$ is $q^2 - 1$. There are $q - 1$ non-zero constants and thus there are $q - 1$ vectors in each equivalence class. The number of classes is

$$n = \frac{q^2 - 1}{q - 1} = q + 1. \tag{19}$$

The number of matrices in $\mathrm{PGL}(2, \mathbb{F}_q)$ is

$$|\mathrm{PGL}(2, \mathbb{F}_q)| = \frac{|\mathrm{GL}(2, \mathbb{F}_q)|}{q - 1}, \tag{20}$$

because there are $q-1$ non-zero constants. The number of matrices in $\mathrm{GL}(2, \mathbb{F}_q)$ is the same as the number of pairs of linearly independent non-zero vectors. The first vector can be chosen in $q^2 - 1$ ways and it determines a set of $q - 1$ linearly dependent vectors and the second vector can be chosen in $(q^2-1)-(q-1) = q^2-q$ ways. Therefore $|\mathrm{GL}(2, \mathbb{F}_q)| = (q^2 - 1)(q^2 - q)$ and

$$|\mathrm{PGL}(2, \mathbb{F}_q)| = (q + 1)q(q - 1) = n(n - 1)(n - 2). \tag{21}$$

The method for obtaining sharply 3-transitive groups given above can be described using a different formalism – the *linear fractional transformation* (also

called *Möbius transformation*):

$$y(x) = \frac{ax + b}{cx + d}, \tag{22}$$

where $a, b, c, d \in \mathbb{F}_q$ and $ad - bc \neq 0$ (otherwise $a/c = b/d = \alpha$ and $y(x) = \alpha$). It acts on the set $\mathbb{F}_q \cup \{\infty\}$ as a permutation. The following conventions regarding the element ∞ are used:

$$y\left(-\frac{d}{c}\right) = \infty, \quad y(\infty) = \begin{cases} \infty & \text{if c=0,} \\ \frac{a}{c} & \text{otherwise.} \end{cases} \tag{23}$$

In fact, the element ∞ corresponds to the vector $\binom{1}{0}$ in the above construction. Note that the inverse of (22) is also a linear fractional transformation:

$$x(y) = \frac{-dy + b}{cy - a}. \tag{24}$$

The same stands for the composition of two linear fractional transformations.

4.6 Sharply 4-transitive and 5-transitive groups

It is known that the *Mathieu group* M_{11} is sharply 4-transitive and therefore the pairwise Hamming distance of distinct elements is at least 8. It consists of $11 \cdot 10 \cdot 9 \cdot 8 = 7920$ elements. It is generated by

$$g_1 = (2, 10)(4, 11)(5, 7)(8, 9)(1)(3)(6), \tag{25}$$
$$g_2 = (1, 4, 3, 8)(2, 5, 6, 9)(7)(10)(11). \tag{26}$$

The *Mathieu group* M_{12} is sharply 5-transitive and the pairwise Hamming distance of distinct elements is also at least 8. It has $12 \cdot 11 \cdot 10 \cdot 9 \cdot 8 = 95040$ elements and is generated by [10]:

$$g_1 = (1, 2)(3, 4)(5, 6)(7, 8)(9, 10)(11, 12), \tag{27}$$
$$g_2 = (1, 2, 3)(4, 5, 7)(8, 9, 11)(6)(10)(12). \tag{28}$$

4.7 Summary of sharply *k*-transitive groups

The Table 1 provides a summary of the sharply k-transitive groups discussed in the Sections 4.1 trough 4.6. These groups match the bound for $G(n, d)$ in Lemma 2 with equality. In Fig. 1 the points of the (n, d)-plane where the maximal value of $G(n, d)$ can be reached are shown. According to Table 1 these points can be divided into 5 infinite classes (indicated with lines) and two *sporadic groups* – the Mathieu groups. It turns out that there are no other groups of maximal size except the Mathieu groups M_{11} and M_{12} between the lines $d \geq 4$ and $d \leq n - 3$:

Theorem 4 (see [6–8, 10]). *A sharply k-transitive group ($k \geq 4$) is isomorphic either to S_n ($n \geq 4$), A_n ($n \geq 6$) or one of the Mathieu groups M_{11} or M_{12}.*

However, the non-existence of maximal groups does not imply that there are no groups with the required properties (see Sect. 6).

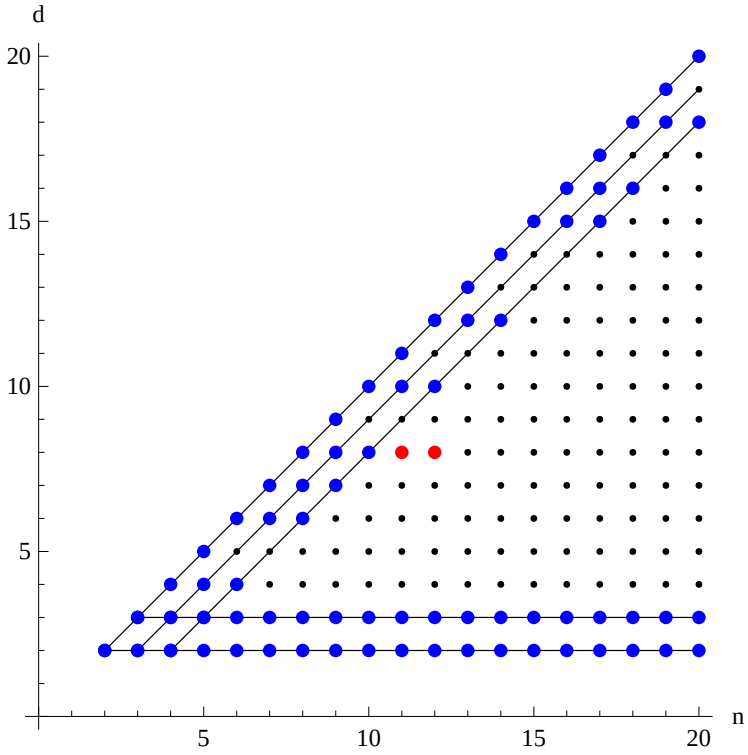

Fig. 1. The maximal permutation groups.

Table 1. The summary of sharply k-transitive groups. The meanings of columns are as follows: d – the pairwise Hamming distance, n – the possible values of the size of the set S (p^m stands for a power of any prime number), $|G_n|$ – the size of the group, G_n – the description of the group.

k	d	n	$\lvert G_n\rvert$	G_n	Section
$n, n-1$	2	any	$n!$	S_n	4.1
$n-2$	3	any	$n!/2$	A_n	4.2
5	8	12	95040	M_{12}	4.6
4	8	11	7920	M_{11}	4.6
3	$n-2$	p^m+1	$n(n-1)(n-2)$	$\mathrm{PGL}(2,\mathbb{F}_{n-1})$	4.5
2	$n-1$	p^m	$n(n-1)$	$y(x)=ax+b$	4.4
1	n	any	n	C_n	4.3

5 Permuting polynomials

The affine transformations considered in Sect. 4.4 were actually linear polynomials of x over the field $\mathbb{F}_n$. The linear fractional transformation considered in Sect. 4.5 is also a polynomial, because for any non-zero $a \in \mathbb{F}_q$ we have $a^{-1} \equiv a^{q-2}$ (this is a consequence of the analog of *Fermat's Little Theorem* $a^{p-1} \equiv 1 \mod p$ for finite fields). From this point of view it is interesting to study the *permuting polynomials* over finite fields, because they can give rise to permutation groups with large pairwise Hamming distance. In this section we give some basic results on groups generated by permuting polynomials.

Let us assume that the size n of the set S on which the permutations act is a power of a prime number (otherwise we can augment the set S with additional elements). Then we can put $\mathrm{S} = \mathbb{F}_n$ and express any permutation (actually any function) f on this set as a polynomial (we assume that f acts trivially on the appended elements if $|\mathrm{S}|$ was not a power of a prime number) as follows:

Definition 7. *The* Lagrange interpolating polynomial *of a function f defined on a finite field $\mathbb{F}_n$ as $f(x_1) = y_1$, $f(x_2) = y_2$, ..., $f(x_n) = y_n$ is given by:*

$$P(x) = \sum_{i=1}^{n} P_i(x), \quad \textit{where} \quad P_i(x) = \prod_{\substack{j=1 \\ j \neq i}}^{n} \frac{x - x_j}{x_i - x_j} y_i. \tag{29}$$

Note that the division is performed in the field $\mathbb{F}_n$ and the denominator always differs from zero.

It is easy to see that the polynomial $P(x)$ mimics the function $f(x)$, i.e., $P(x) = f(x)$ for all $x \in \mathbb{F}_n$. As an example we will consider the permutation groups over the field $\mathbb{F}_5$.

The symmetric group S_5 is generated by $g_1(x) = x + 1$ and $g_2(x) = x^3$. It consists of all polynomials of the form:

$$ax^3 + bx^2 + 2a^3b^2x + d, \quad a \neq 0, \tag{30}$$

$$cx + d, \quad c \neq 0. \tag{31}$$

The alternating group A_5 is generated by $g_1(x) = x + 1$ and $g_2(x) = 2x^3$. It consists of all polynomials of the form (30) and (31), except that in addition we require that $a \in \{2, 3\}$ (non-squares) and $c \in \{1, 4\}$ (squares) respectively. Affine transformations are generated by $g_1(x) = x + 1$ and $g_2(x) = 2x$ and they are of the form (31). The cyclic group C_5 is generated by $g_1(x) = x + 1$ and is of the form (31) where $c = 1$.

As an example of a permutation group generated by polynomials when $|S|$ is not a power of a prime number, we can mention the case $G(6, 4) = 120$. The corresponding group is generated by $g_1(x) = 6x + 5$ and $g_2(x) = x^4 + 3x + 1$ where both polynomials are modulo 7.

Table 2. Experimentally obtained results for $G(n, d)$. The columns have the following meaning: n – the size of the set S, d – the pairwise Hamming distance, $G(n, d)$ – the size of the group obtained, "Bound" – the upper bound for $G(n, d)$ according to Lemma 2, "Generators" – the two generators of the group.

n	d	$G(n, d)$	Bound	Generators
7	4	168	840	$6, 4, 3, 2, 5, 1, 7$ $6, 1, 7, 5, 2, 3, 4$
8	5	336	1680	$3, 8, 6, 2, 4, 5, 1, 7$ $7, 4, 6, 3, 1, 5, 2, 8$
8	4	1344	6720	$2, 6, 8, 4, 5, 7, 1, 3$ $7, 4, 3, 5, 1, 8, 6, 2$
9	6	1512	3024	$4, 5, 1, 8, 3, 7, 6, 2, 9$ $3, 4, 8, 5, 7, 1, 6, 9, 2$
9	5	1512	15120	$9, 4, 1, 6, 5, 2, 7, 8, 3$ $1, 4, 5, 3, 7, 9, 8, 2, 6$
9	4	1512	60480	$7, 2, 8, 3, 5, 6, 9, 4, 1$ $6, 1, 3, 8, 2, 4, 9, 5, 7$
10	7	720	5040	$3, 9, 5, 7, 4, 8, 10, 6, 1, 2$ $7, 9, 4, 5, 3, 6, 8, 1, 10, 2$
10	6	1512	30240	$8, 2, 10, 7, 4, 3, 1, 6, 5, 9$ $1, 2, 8, 5, 10, 6, 3, 7, 9, 4$
10	5	1512	151200	$1, 10, 3, 9, 6, 8, 5, 4, 7, 2$ $1, 10, 8, 3, 2, 4, 5, 7, 6, 9$
10	4	1920	604800	$5, 1, 4, 8, 9, 7, 6, 10, 2, 3$ $7, 8, 2, 1, 10, 3, 9, 6, 4, 5$
15	12	2520	32760	$7, 2, 4, 5, 11, 10, 13, 15, 3, 9, 6, 8, 14, 12, 1$ $9, 15, 11, 6, 4, 2, 10, 13, 7, 12, 8, 1, 14, 3, 5$
16	12	40320	524160	$16, 5, 6, 12, 14, 13, 11, 1, 10, 3, 7, 4, 15, 8, 9, 2$ $6, 7, 14, 8, 15, 3, 12, 2, 9, 10, 13, 11, 4, 16, 1, 5$

6 Experimental results

We performed computer experiments to find permutation groups with pairwise Hamming distance in the region between $d \geq 4$ and $d \leq n-3$. The obtained results for $n = 7, 8, 9, 10$ are shown in Table 2. In addition we mention also two large groups for $n = 15$ and $n = 16$.

These groups were obtained by choosing two random permutations g_1 and g_2 and computing their closure with respect to the product of permutations. If at some point the distance between any two distinct obtained permutations became less than some predefined d_{min}, the process was terminated and restarted with another random generators g_1 and g_2. Some of the groups obtained in this way have very interesting properties:

(1) $G(7, 4)$ has $168 = 7 \cdot 6 \cdot 4$ elements and is isomorphic to the automorphism group of the *Fano plane*.
(2) $G(8, 4)$ has $1344 = 8 \cdot 168 = 8 \cdot 7 \cdot 6 \cdot 4$ elements. This group has the property, that the stabilizers of any element form a group that is isomorphic to the automorphism group of the Fano plane. This group also has a property that for any 3-tuples x and y of distinct elements there are exactly 4 permutations that send x to y. It is isomorphic to the automorphism group of the *octonion* multiplication table.
(3) $G(9, 6)$ has $1512 = 9 \cdot 168 = 9 \cdot 8 \cdot 7 \cdot 3$ elements and it has the same stabilizer property, but for each 3-tuples x and y there are exactly 3 permutations that send x to y.
(4) $G(15, 12)$ has $2520 = 15 \cdot 168 = 15 \cdot 14 \cdot 12$ elements and it also has the stabilizer property, but for each 2-tuples x and y there are exactly 12 permutations that send x to y.
(5) $G(16, 12)$ has $40320 = 16 \cdot 15 \cdot 168 = 16 \cdot 15 \cdot 14 \cdot 12$ elements. The stabilizers of any two elements form a group that is isomorphic to the automorphism group of the Fano plane. For any 3-tuples x and y there are exactly 12 permutations that send x to y.

References

1. A.Ambainis, R.Freivalds. 1-way quantum finite automata: strengths, weaknesses and generalizations. *Proc. IEEE FOCS'98*, pp. 332–341, 1998.
2. A.Ambainis. The complexity of probabilistic versus deterministic finite automata. *Lecture Notes in Computer Science,* Springer, v.1178, pp.233–237, 1996.
3. R.Freivalds. Non-constructive methods for finite probabilistic automata. Accepted for *The 11th International Conference Developments in Language Theory DLT-2007*, Turku, Finland, July 3-5, 2007.
4. D.Aharonov, A.Kitaev, N.Nisan. Quantum circuits with mixed states. *Proc. STOC 1998*, pp. 20-30, 1998.
5. A.Ambainis, M.Beaudry, M.Golovkins, A.Ķikusts, M.Mercer, D.Thérien. Algebraic Results on Quantum Automata. *Theory Comput. Syst.*, v. 39(1), pp. 165–188, 2006.
6. A.Wool. Sharply Transitive Permutation Groups, 2006.

7. P.J.Cameron. *Permutation Groups*, Chapter 12, pp. 611–645 in R.L.Graham, M.Grötschel, L.Lovàsz. *Handbook of Combinatorics*, Vol 1, Elsevier Science B.V., The MIT Press, 1995.
8. G.F.Pilz. Near-rings and Near-fields, pp. 463–498 in M.Hazewinkel. *Handbook of Algebra*, Vol 1, Elsevier Science B.V., 1996.
9. K.Itô. *Encyclopedic Dictionary of Mathematics*, 2nd ed., 151 H. *Transitive Permutation Groups*, pp. 591–593, The MIT Press, 1987.
10. R.F.Bailey. Decoding the Mathieu Group M_{12}.
11. P.J.Cameron. Permutation Codes, talk at CGCS, 2007.
12. E.Artin. Beweis des allgemeinen Reziprozitätsgesetzes. *Mat. Sem. Univ. Hamburg*, B.5 , 353–363, 1927.
13. M.Aschbacher. *Finite Group Theory*, (Cambridge Studies in Advanced Mathematics), Cambridge University Press, 2nd edition, 2000.
14. E.Bach, J.Shallit. *Algorithmic Number Theory.*, vol. 1, MIT Press, 1996.
15. A.Cobham. The recognition problem for the set of perfect squares. *Proc. 7th Ann. Symp. Switching and Automata Theory*, pp. 78–87, 1966.
16. Y.L.Ershov. Theory of numberings, *Handbook of computability theory* (E.R.Griffor, ed.), North-Holland, Amsterdam, pp. 473-503, 1999.
17. R.Freivalds. On the growth of the number of states in result of the determinization of probabilistic finite automata. *Avtomatika i Vichislitel'naya Tekhnika*, No. 3, pp. 39–42, 1982. (Russian)
18. N.Z.Gabbasov, T.A.Murtazina. Improving the estimate of Rabin's reduction theorem. *Algorithms and Automata*, Kazan University, pp. 7–10, 1979. (Russian)
19. P.Garret. *The Mathematics of Coding Theory.*, Pearson Prentice Hall, Upper Saddle River, 2004.
20. C.Hooley. On Artin's conjecture. *J.ReineAngew.Math.*, v. 225, pp. 229–220, 1967.
21. D.R.Heath-Brown. Artin's conjecture for primitive roots. *Quart. J. Math. Oxford*, v. 37, pp. 27–38, 1986.
22. A.Kondacs, J.Watrous. On the power of quantum finite state automata. *Proc. IEEE FOCS'97*, pp. 66–75, 1997.
23. C.Moore, J.Crutchfield. Quantum automata and quantum grammars. *Theor. Comput. Sci.*, v. 237(1-2), pp. 275–306, 2000.