Group Actions and Elementary Number Theory*

D. P. Patil[1] AND U. Storch[2]

Abstract | In this article the shuffling of cards is studied by using the concept of a group action. We use some fundamental results in Elementary Number Theory to obtain formulas for the orders of some special shufflings, namely the Faro and Monge shufflings and give necessary and sufficient conditions for the Monge shuffling to be a cycle. In the final section we extend the considerations to the shuffling of multisets.

Introduction

In this (expository) article the shuffling of cards is used as a motivation for introducing the concept of a group action. It combines the concept of an abstract group with the original idea of a group as a transformation group and is fundamental in almost all parts of mathematics and its applications in physical sciences. In Section 1 we discuss besides basic properties and constructions several examples from various branches of mathematics demonstrating the ubiquity of group operations. We want to demonstrate that the use of group actions very often clarifies and simplifies the modeling in mathematics and other domains and may lead to new and interesting questions. The applications to some special shuffling methods in Sections 3 and 4 require certain fundamental results in Elementary Number Theory. In order to make this article self-contained, in Section 2 we give a rather thorough exposition (including proofs) on Elementary Number Theory. The main topics there are (General) Chinese Remainder Theorem, Prime Residue Class Groups and their structure, Quadratic Residues and (some examples of) Diophantine Equations. The exponentiation maps in groups including the discrete logarithm problems are also discussed. Results in Section 1 and Section 2 are standard but our approach to some of them is different and many of them are not found together in a single reference. In Section 3 we deal with some special shufflings, namely the Faro and Monge shufflings, again with many examples. In Theorem 3.4 and Theorem 3.6 we obtain formulas for the orders of the Faro and Monge shufflings, respectively, which are partially already mentioned in the books [1] and [2]. In Corollary 3.9 we give necessary and sufficient conditions for the Monge shuffling to be a cycle. In Section 4 we consider the shuffling of multisets which leads canonically to the study of cosets and double cosets as elements of the respective orbit spaces.

[1]Department of Mathematics, Indian Institute of Science, Bangalore 560 012, India

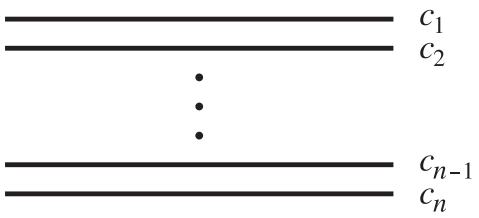
[2]Fakultät für Mathematik, Ruhr-Universität Bochum, D-44780 Bochum, Germany

E-mail: [1]patil@math.iisc.ernet.in , [2]Uwe.Storch@ruhr-uni-bochum.de

* This article is an extended version of a lecture entitled "Applications of Number Theory – Shuffling of Cards" given by the second author on invitation by Professor Dr. R. Ravindran at the Atria Institute of Technology, Bangalore, on February 23, 2008, while visiting the Department of Mathematics, Indian Institute of Science, Bangalore 560 012. Both authors thank DAAD for financial support.

§ 1 Operations of Groups – Shuffling and other Examples

1.1 Shuffling Let C denote a set of n playing cards, $n \in \mathbb{N}^* := \{1,2,\dots\}$. From this set one forms a pack or a stack by arranging the cards of C in a sequence $(c_1,\dots,c_n)$ in which every card occurs exactly once. In other words a pack is a *bijective* map $\mathbf{c}:[1,n]\to C$ from $[1,n]:=\{1,\dots,n\}$ onto C. We set $c_i := \mathbf{c}(i)$ for $i=1,\dots,n$ and write $\mathbf{c}:=(c_1,\dots,c_n)$. We illustrate such a pack $\mathbf{c}$ by a pile with c_1 as the top card and c_n as the bottom card.



The set of all packs is denoted by $\mathfrak{P}=\mathfrak{P}_C$. It is a finite set of cardinality $\#\mathfrak{P}=n!=(\#C)!$. It simplifies the investigation to distinguish carefully the set C of cards and the set $[1,n]$ of the possible positions of a card in a stack. One reason for this is that a priori no fixed order for C is given. Even, if there is such a canonical order then it depends on the rules of the game you are playing.

A shuffling is a re-arrangement of the cards in a pack $\mathbf{c}=(c_1,\dots,c_n)\in\mathfrak{P}_C$. Such a shuffling can be described with a permutation of the set $[1,n]$, i. e. with an element of the permutation group $\mathfrak{S}([1,n])=\mathfrak{S}_n$, under the following two view points:

(1) The shuffling brings the card c_i from the i-th place to the $\sigma(i)$-th place, $i=1,\dots,n$. This defines a permutation $\sigma\in\mathfrak{S}_n$.

(2) After shuffling the card at the i-th place is the card $c_{\tau(i)}$, $i=1,\dots,n$. This defines also a permutation $\tau\in\mathfrak{S}_n$.

In case (1) the new pack is $\mathbf{c}\,\sigma^{-1}=\mathbf{c}\circ\sigma^{-1}$ and in case (2) it is $\mathbf{c}\,\tau=\mathbf{c}\circ\tau$.[1] Since $\mathbf{c}:[1,n]\to C$ is also bijective, it follows that $\sigma^{-1}=\tau$ or $\sigma=\tau^{-1}$. We have to decide and do this in such a way that *we consider a shuffling always under view point* (1) and call

$$\sigma * \mathbf{c} := \mathbf{c}\,\sigma^{-1}$$

the stack obtained from $\mathbf{c}\in\mathfrak{P}_C$ by shuffling with the permutation $\sigma\in\mathfrak{S}_n$.

[1] We emphasize that we always perform the compositions of maps from right to left and this is also done for permutations of a set. Therefore, for $\varphi,\psi\in\mathfrak{S}_n$, in $\varphi\psi=\varphi\circ\psi$ first we apply ψ and then φ, for example, $\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}\circ\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}=\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}$ and $\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}\circ\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}=\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}$.

1.1.1 Example From a given pack $\mathbf{c}$ by shifting the bottom card c_n of $\mathbf{c}$ to the top the new stack is $\gamma_n * \mathbf{c} = \mathbf{c}\,\gamma_n^{-1}$ with $\gamma_n := \begin{pmatrix} 1 & 2 & \dots & n-1 & n \\ 2 & 3 & \dots & n & 1 \end{pmatrix}$. On the other hand, the pack $\gamma_n^{-1} * \mathbf{c} = \mathbf{c}\,\gamma_n$ is obtained from $\mathbf{c}$ by shifting the top card c_1 to the bottom which is different from the former one if $n > 2$. Note that $\gamma_n^{-1} = \begin{pmatrix} 1 & 2 & \dots & n-1 & n \\ n & 1 & \dots & n-2 & n-1 \end{pmatrix}$.

Altogether, we get a map

$$\mathfrak{S}_n \times \mathfrak{P}_C \to \mathfrak{P}_C\,, (\sigma, \mathbf{c}) \mapsto \sigma * \mathbf{c} := \mathbf{c}\,\sigma^{-1}\,,$$

which is a simply transitive operation of the permutation group $\mathfrak{S}_n$ on the set $\mathfrak{P}_C$ of stacks of C. We shall explain these concepts in the next subsection.

1.2 Operations of groups on sets Let us recall the important concept of operation of an (arbitrary) group on an (arbitrary) set. Let G be a (multiplicatively written) group with the identity element $e := e_G$. An operation or an action of G (as a group) on a set X [2] is a map $G \times X \to X\,, (g,x) \mapsto gx$, with the following properties:

(1) $(gh)x = g(hx)$ for all $g, h \in G$ and all $x \in X$.
(2) $ex = x$ for all $x \in X$.

More precisely, an operation of G on X as described above is an operation from the left or a left operation. An operation from the right or a right operation is a map $X \times G \to G, (x,g) \mapsto xg$, with $x(gh) = (xg)h$ and $xe = x$ for all $g, h \in G$ and all $x \in X$. If $(x,g) \mapsto xg$ is an operation from the right then $(g,x) \mapsto gx := xg^{-1}$ is an operation from the left, and conversely. Therefore, in principle, left and right operations are interchangeable.

Let $G \times X \to X$ be an operation (from the left). The operation of g on X is the map $\vartheta_g : X \to X\,, x \mapsto gx$. Then, the above conditions (1) and (2) are equivalent to $(1')$ $\vartheta_{gh} = \vartheta_g \vartheta_h$ and $(2')$ $\vartheta_e = \mathrm{id}_X$. In particular, ϑ_g is a permutation of X with inverse $(\vartheta_g)^{-1} = \vartheta_{g^{-1}}$. Therefore the map $\vartheta : G \to \mathfrak{S}(X)\,, \vartheta(g) := \vartheta_g$, is a group homomorphism. Conversely, if $\vartheta : G \to \mathfrak{S}(X)$ is a group homomorphism, then the map $G \times X \to X$ defined by $(g,x) \mapsto gx := \vartheta(g)(x)$ is an operation of G on X.

A set X with an operation of a group G (from the left) is called a G-set or a G-space and the group homomorphism $\vartheta : G \to \mathfrak{S}(X)$ belonging to it is called the action homomorphism of the G-set X.[3]

The kernel of ϑ, i. e. the set of $g \in G$ with $\vartheta_g = \mathrm{id}_X$ or with $gx = x$ for all $x \in X$, is called the kernel of the operation. If this kernel is trivial, then the action is called faithful or effective. If this kernel is the whole group G, i. e. if $gx = x$ for all $g \in G$ and all $x \in X$, then the action is the so-called trivial action.

[2] The study of sets with group operation was initiated by Felix Klein in his famous Erlanger program "Vergleichende Betrachtungen über neuere geometrische Forschungen" from 1872, whereby Klein considered only transformation groups (see also Footnote 5), especially Lie groups occurring as transformation groups. His main examples were derived from projective groups and their subgroups operating canonically on projective spaces. The more general concept we introduce here goes back basically to Hermann Weyl. Some special cases occurred already in 1854 in the work of Arthur Cayley on abstract group theory.

[3] For a right operation $X \times G \to X$ the action homomorphism $g \mapsto (\eta_g : x \mapsto xg)$ is an anti-homomorphism $\eta : G \to \mathfrak{S}(X)$ of groups: $\eta(gh) = \eta_{gh} = \eta_h \circ \eta_g = \eta(h)\eta(g)$ for all $g, h \in G$, i. e. a homomorphism $G^{\mathrm{op}} \to \mathfrak{S}(X)$ from the opposite group G^{op} (with binary operation $g \diamond h := hg$) in $\mathfrak{S}(X)$. Therefore, *a right operation of G is the same as a left operation of G^{op}* (and conversely).

The operation of G on X defines in a natural way an equivalence relation $\sim_G$ on X. Elements $x, y \in X$ are related if y is obtained from x by the operation ϑ_g of a suitable element $g \in G$, i. e.

$$x \sim_G y \iff \text{there exists } g \in G \text{ with } y = gx.$$

That this is indeed an equivalence relation is easily checked by using the conditions (1) and (2) for a group operation.

The equivalence class

$$Gx := \{gx \mid g \in G\}$$

of an element $x \in X$ is called the (G-) orbit of x. The orbit space of X (with respect to the given operation), i. e. the set of all orbits Gx, $x \in X$, is denoted by

$$X \backslash G.$$

Many authors denote the quotient set $X \backslash G$ by $G \backslash X$. For a right operation the set of all orbits $xG := \{xg \mid g \in G\}$, $x \in X$, is denoted by X/G.

To understand an orbit Gx one considers the isotropy group or the stabilizer of $x \in X$. This is the subgroup $G_x := \{g \in G \mid gx = x\}$ of those elements $g \in G$ for which x is a fixed point of ϑ_g. The point $x \in X$ is a fixed point of the operation if and only if $G_x = G$. The set of all fixed points is denoted by

$$\mathrm{Fix}_G X \qquad (\text{or } X^G).$$

Obviously, in any case the fibres of the canonical surjective map $f_x : G \to Gx$, $g \mapsto gx$, are the left-cosets of G_x in G, i. e.

$$f_x^{-1}(gx) = \{h \in G \mid hx = gx\} = \{h \in G \mid g^{-1}hx = x\} = gG_x.$$

This proves the following important theorem, which is used very often:

1.3 Theorem (Orbit-Stabilizer Theorem) *Let X be a G-set. The cardinality $\#Gx$ of the orbit Gx of x is the index $[G : G_x] := \#(G/G_x)$ of the stabilizer G_x of x in G,* i.e.

$$\#Gx = [G : G_x].$$

In particular, if G is finite, then the cardinality $\#Gx$ of Gx divides the order $\#G$ of G. – Furthermore, the stabilizers of the elements in the same orbit are conjugate subgroups, more precisely,

$$G_{gx} = g\,G_x\,g^{-1}, \quad g \in G, x \in X.$$

We remark that $\#Gx$ divides $\#G$ in case G is finite follows from the general equality

$$[G : H] \cdot \#H = \#G$$

for any subgroup H of a finite group G which again is a consequence of the following: *All cosets gH have the same cardinality as H*, since the maps $h \mapsto gh$ from H to gH are bijective. Furthermore, one obtains Lagrange's Theorem: *The order of a subgroup of a finite group divides the order of the group.*[4]

[4]While Lagrange did not have the group concept – not even that of a group of permutations – he was the first to realize the significance of the study of permutations of the roots for the theory of equations. Moreover, his work on the theory of equations published in 1770 stimulated the later work of Cauchy and Galois and contained in essence the proof of what we call now Lagrange's Theorem.

If X is finite and if we count the elements of X with the help of orbits of a G-operation on X, then we get the so-called class equation:

1.4 Theorem (Class Equation) *Let G be a finite group and let X be a finite G-set. Then*

$$\#X = \sum_{Gx \in X\backslash G} \#Gx = \sum_{Gx \in X\backslash G} [G : G_x] = \#\mathrm{Fix}_G X + \sum_{\substack{Gx \in X\backslash G, \\ Gx \neq \{x\}}} [G : G_x]\,.$$

A group operation $G \times X \to X$ is called transitive if it has exactly one orbit, i. e. if $X \neq \emptyset$ and if $Gx = X$ for one $x \in X$ and hence for all $x \in X$. Equivalently, G operates transitively on X if $X \neq \emptyset$ and if for arbitrary $x, y \in X$ there exists $g \in G$ with $y = gx$.

The last assertion in Theorem 1.3 implies that *the stabilizers G_x of the elements $x \in X$ with respect to a transitive operation of G on X form a full conjugacy class of subgroups of G.*

A G-space with a transitive operation of the group G is called a homogeneous G-space, see also Example 1.10.

A group operation $G \times X \to X$ is called simply transitive if it is transitive and if one and hence all stabilizers G_x, $x \in X$, are trivial. Equivalently, G operates simply transitively on X if $X \neq \emptyset$ and if for arbitrary $x, y \in X$ there exists exactly one $g \in G$ with $y = gx$.

If all isotropy groups are trivial, i. e. if G operates simply transitively on each orbit of the operation, then the operation is called free.

1.5 Examples Let G be a group.

(1) (Left regular or the Cayley operation) The multipication $G \times G \to G$ in the group G is the most natural simply transitive operation of G onto itself. The corresponding action homomorphism $\vartheta : G \to \mathfrak{S}(G)$ maps g to the left multiplication $\mathrm{L}_g : G \to G$, $x \mapsto gx$. This operation is called the (left) regular operation or the Cayley operation of G on itself.

(2) (Conjugation operation) A somewhat less canonical example of an operation of G onto itself is the conjugation operation: $(g,x) \mapsto gxg^{-1}$, $g \in G$, $x \in G$. The corresponding action homomorphism is $\mathrm{K} : G \to \mathrm{Aut}\, G \subseteq \mathfrak{S}(G)$, $g \mapsto \mathrm{K}_g \in \mathrm{Aut}\, G$, where $\mathrm{K}_g : G \to G$ is the inner automorphism $x \mapsto gxg^{-1}$ of G by g. The orbits of this operation on G are called the conjugacy classes in G and the fixed point set $\mathrm{Fix}_G\, G$ is the center $\mathrm{Z}(G) := \{x \in G \mid gx = xg \text{ for all } g \in G\}$ of G, which is also the kernel of this action, i. e. $\ker \mathrm{K} = \mathrm{Z}(G)$. Moreover, if G is finite, then the class equation of G is:

$$\#G = \#\mathrm{Z}(G) + \sum_{i=1}^{r} \#C_i\,,$$

where $C_1, \ldots, C_r$ denote the distinct conjugacy classes with cardinality > 1. If $x_i \in C_i$, then $\#C_i = [G : \mathrm{C}_G(x_i)]$, where, for $x \in G$, $\mathrm{C}_G(x) = \{g \in G \mid gx = xg\}$ is the subgroup of those elements of G which commute with x; it is called the centralizer of x in G. Note that the numbers $\#\mathrm{Z}(G)$ and $\#C_i$, $i = 1, \ldots, r$, are all divisors of the order $\#G$ of the group. The number of all conjugacy classes, i. e. $\#\mathrm{Z}(G) + r$ is called the class number of G.

As an application we note the following: *If G is a non-trivial finite p-group,* i. e. $\#G = p^m$ *with p prime and $m \in \mathbb{N}^*$, then G has a non-trivial center.* For a proof note that, since $\#G$ is

a power of a prime number p, in the above class equation of G of the order $\#G$ as well as all other terms $\#C_i$ are divisible by p and hence p divides $\#\mathrm{Z}(G)$.

More generally, from the class equation in 1.4 it follows that: *If a finite p-group G operates on a finite set X, then the congruence* $\#X \equiv \#\operatorname{Fix}_G X \bmod p$ *holds.*

(3) (Cauchy's Theorem) Let G be a finite group of order n and let p be a prime number. On the set G^p of p-tuples of G, the cyclic group $\mathbf{Z}_p$ operates by $(a,(x_1,\ldots,x_p)) \mapsto (x_{1+a},\ldots x_{p+a})$, where the sum with a and the indices $1,\ldots,p$ is the addition in the group $\mathbf{Z}_p$. The fixed points of this operation are the constant tuples $(x,\ldots,x)$. If $x_1\cdots x_p = (x_1\cdots x_r)(x_{r+1}\cdots x_p) = e$, we also have $(x_{r+1}\cdots x_p)(x_1\cdots x_r) = e$ for all $r = 1,\ldots,p-1$ and hence the subset $X := \{(x_1,\ldots,x_p) \in G^p \mid x_1\cdots x_p = e\}$ of G^p is $\mathbf{Z}_p$-invariant, From the class equation of the $\mathbf{Z}_p$-set X, we get $n^{p-1} = \#X \equiv \#\operatorname{Fix}_{\mathbf{Z}_p} X \bmod p$. Therefore, if p divides n, then p also divides $\#\operatorname{Fix}_{\mathbf{Z}_p} X$. In particular, there exists $x \in G$, $x \neq e$ such that $x^p = e$. This proves the following well-known Theorem of Cauchy: *A finite group G contains an element of order p for every prime divisor p of* $\#G$. Furthermore, if p does not divide n, then $\operatorname{Fix}_{\mathbf{Z}_p} X = \{(e,e,\ldots,e)\}$ by Lagrange's Theorem and hence the well-known congruence $n^{p-1} \equiv 1 \bmod p$ (Fermat's Little Theorem).

Maps between G-sets which are compatible with the operation of G are called G-homomorphisms. More precisely: Let X and Y be G-sets with the operations $G\times X \to X$ and $G\times Y \to Y$. Then a G-homomorphism or just a G-morphism from X to Y is a map $f : X \to Y$ with $f(gx) = gf(x)$ for all $x \in X$ and $g \in G$, i. e. the diagram

$$\begin{array}{ccc} G\times X & \longrightarrow & X \\ {\scriptstyle \mathrm{id}\times f}\downarrow & & \downarrow{\scriptstyle f} \\ G\times Y & \longrightarrow & Y \end{array}$$

is commutative. Such a G-morphism induces the map $\overline{f} : X/G \to Y/G$, $\overline{f}(Gx) = Gf(x)$, $x \in X$, of the corresponding orbit spaces. The concept of G-morphism can be generalized in the following way: Let $\varphi : G \to G'$ be a group homomorphism. Furthermore, let X and X' be a G-set and a G'-set respectively. Then a map $f : X \to X'$ is called a φ-morphism from X to X' if the diagram

$$\begin{array}{ccc} G\times X & \longrightarrow & X \\ {\scriptstyle \varphi\times f}\downarrow & & \downarrow{\scriptstyle f} \\ G'\times X' & \longrightarrow & X' \end{array}$$

is commutative, i. e. if $f(gx) = \varphi(g)f(x)$ for all $x \in X$, $g \in G$.

By an isomorphism between a G-set X and a G'-set X', we always mean a bijective map $f : X \to X'$ which is a φ-morphism for some group isomorphism $\varphi : G \to G'$. In this case, the diagram

$$\begin{array}{ccc} G & \xrightarrow{\vartheta} & \mathfrak{S}(X) \\ {\scriptstyle \varphi}\downarrow & & \downarrow{\scriptstyle \mathrm{K}_f} \\ G' & \xrightarrow{\vartheta'} & \mathfrak{S}(X') \end{array}$$

is commutative, i. e. $\vartheta' = \mathrm{K}_f\,\vartheta\,\varphi^{-1}$, where $\mathrm{K}_f : \mathfrak{S}(X) \to \mathfrak{S}(X')$ denote the "conjugation" $\sigma \mapsto f\sigma f^{-1}$.

We illustrate the above concepts by some examples from various branches of mathematics. Group operations allow to handle entities of very different mathematical origins in a flexible way, while retaining essential structural aspects of many objects in abstract algebra and beyond.

1.6 Examples (Induced operations) From given group operations new operations can be constructed in many ways.

(1) An operation $G \times X \to X$ of the group G on a set X can be restricted to every subgroup $H \subseteq G$. An element $h \in H$ operates in the same way as the given operation as an element of G. The operation map $H \times X \to X$ is the restriction of the given operation map $G \times X \to X$ to $H \times X$, and the isotropy group is the intersection of H with the isotropy group G_x of x, i. e. $H_x = H \cap G_x$. The action homomorphism $H \to \mathfrak{S}(X)$ is simply the restriction of the given action homomorphism $\vartheta : G \to \mathfrak{S}(X)$. More generally, if $\varphi : F \to G$ is a group homomorphism, then the given operation of G on X induces an operation of F on X by $fx := \varphi(f)x$ for $f \in F$, $x \in X$, with action homomorphism $\vartheta \circ \varphi : F \to \mathfrak{S}(X)$. In principle, every group action can be obtained in this way by the canonical action $\mathfrak{S}(X) \times X \to X$, $\sigma x := \sigma(x)$, $\sigma \in \mathfrak{S}(X), x \in X$, of the full permutation group $\mathfrak{S}(X)$ on X. The action homomorphism $\vartheta : G \to \mathfrak{S}(X)$ induces the given operation $G \times X \to X$. The identity of X is a ϑ-morphism. The trivial homomorphism $\varepsilon : G \to \mathfrak{S}(X)$ (with $\varepsilon(g) = \mathrm{id}_X$) gives the trivial action of G on X. For any subgroup $G \subseteq \mathfrak{S}(X)$, G operates faithfully on X by restricting the canonical operation $\mathfrak{S}(X) \times X \to X$. If not mentioned otherwise we consider this canonical operation for every given subgroup $G \subseteq \mathfrak{S}(X)$. If X is finite then $G \subseteq \mathfrak{S}(X)$ operates transitively on X if and only if $G \cap \mathfrak{S}(X \setminus \{x\})$ has index $\#X$ in G for one (and hence all) $x \in X$.

Any faithful operation can be identified with the canonical operation of a subgroup of a permutation group. An arbitrary operation of G with action homomorphism ϑ induces canonically a faithful operation of $G/\ker\vartheta$. Note that the kernel $\ker\vartheta$ of the operation is the intersection of all stabilizers $G_x, x \in X$.[5]

(2) A given group operation $G \times X \to X$ can be restricted to G-invariant subsets. A subset $Y \subseteq X$ is called G-invariant if $gy \in Y$ for all $g \in G$ and all $y \in Y$. For such a G-invariant subset $Y \subseteq X$ the operation map $G \times X \to X$ maps $G \times Y$ into Y and defines an operation $G \times Y \to Y$ of G on Y. Besides the empty set the smallest G-invariant subsets are the G-orbits. Therefore, a subset $Y \subseteq X$ is G-invariant if and only if it is a union of G-orbits.

(3) An operation $G \times X \to X$ of G on X induces an operation of G on the power set $\mathfrak{P}(X)$ of X by $gA := \{gx \mid x \in A\}, g \in G, A \in \mathfrak{P}(X)$. The set $\mathfrak{P}_\alpha(X)$ of subsets of X of a given cardinality α (which is also denoted by $\binom{X}{\alpha}$) is a G-invariant subset of $\mathfrak{P}(X)$. If X is finite, then the $\mathfrak{P}_\alpha(X), \alpha \leq \#X$, are the orbits of the canonical operation of $\mathfrak{S}(X)$ on $\mathfrak{P}(X)$ (induced by the canonical operation of $\mathfrak{S}(X)$ on X). If X is infinite, then the subsets $\mathfrak{P}_\alpha(X), \alpha < \#X$, are again full orbits of the canonical operation of $\mathfrak{S}(X)$ on $\mathfrak{P}(X)$ (because of the well-known equality $\alpha + \beta = \mathrm{Max}(\alpha, \beta)$ for cardinalities α, β with β infinite, see for instance [5, Teil 1, § 7]), but $\mathfrak{P}_{\#X}(X)$ decomposes into more than one orbit. (For each cardinality $\beta \leq \#X$, there is exactly one such orbit, namely the set $\{Y \subseteq X \mid \#Y = \#X, \#(X \setminus Y) = \beta\}$.) In general, the stabilizer of a subset $Y \subseteq X$ is the group of permutations $\sigma \in \mathfrak{S}(X)$ with $\sigma(Y) = Y$. Then $\sigma(X \setminus Y) = X \setminus Y$ too, and the stabilizer is the product group $\mathfrak{S}(Y) \times \mathfrak{S}(X \setminus Y)$ identified canonically with a subgroup of $\mathfrak{S}(X)$. In particular, for a finite set X with $n := \#X$ and a

[5]Historically, by definition groups were transformation groups, i. e. subgroups of permutation groups of sets with their canonical operations. In particular, symmetry groups were considered as such transformation groups operating on the structures under consideration. Such symmetry groups, particularly the continuous Lie groups, play an important role in many academic disciplines, for example, can be used to understand fundamental physical laws underlying special relativity and symmetry phenomena in molecular chemistry. – The axioms for group operations are derived from these concrete examples. Abstract groups were introduced only in the late 19th century, for instance by Dyck (1882) and Weber (1882, 1895).

subset $Y \subseteq X$ with $\#Y := m \leq n$, it follows from Theorem 1.3 that

$$\binom{n}{m} := \#\mathfrak{P}_m(X) = \#\mathfrak{S}(X)Y = [\mathfrak{S}(X) : (\mathfrak{S}(Y) \times \mathfrak{S}(X \setminus Y))] = \frac{n!}{m!(n-m)!}.$$

More generally, if X is a G-set and Y is an H-set, then the set $\mathrm{Map}(X,Y) = Y^X$ of all maps from X to Y is a $G \times H$-set in a canonical way: $(g,h)f(x) = hf(g^{-1}x), (g,h) \in G \times H, f \in Y^X, x \in X$, i. e. $\vartheta_{(g,h)}f = \vartheta_h f \vartheta_{g^{-1}}$. The operation of G on $\mathfrak{P}(X)$ as described above is obtained by identifying $\mathfrak{P}(X)$ with $\{0,1\}^X$ and adding the (trivial) operation of the trivial group $H = \{e_H\}$ on $\{0,1\}$.

The set of bijective maps $X \to Y$ is an $G \times H$- invariant subset of Y^X and therefore carries an induced $G \times H$-operation. In particular, if $X = \{1,\ldots,n\}$ and $Y = C$ is a set of cards with n elements, then the set $\mathfrak{P}_C$ of packs carries the canonical $(\mathfrak{S}(X) = \mathfrak{S}_n)$-operation induced by the canonical operation of $\mathfrak{S}_n$ and the operation of the trivial group $\{\mathrm{id}_C\}$ on C. This is exactly the operation we introduced in Subsection 1.1. But, on C the full permutation group $\mathfrak{S}(C)$ also operates canonically. Together with the canonical operation of $\mathfrak{S}_n$ on $\{1,\ldots,n\}$ this gives the operation

$$(\rho,\sigma) * \mathbf{c} = \rho\,\mathbf{c}\,\sigma^{-1}$$

of $\mathfrak{S}(C) \times \mathfrak{S}_n$ on $\mathfrak{P}_C$ which extends the above operation of $\mathfrak{S}_n = \{\mathrm{id}_C\} \times \mathfrak{S}_n$ on $\mathfrak{P}_C$ and which will also be used later. The operation of $\mathfrak{S}_n = \{\mathrm{id}\} \times \mathfrak{S}_n$ and of $\mathfrak{S}(C) = \mathfrak{S}(C) \times \{\mathrm{id}\}$ are obviously simply transitive. For $n > 1$, the operation of the product $\mathfrak{S}(C) \times \mathfrak{S}_n$ is only transitive but not simply transitive. For a pack $\mathbf{c} : [1,n] \to C$ the stabilizer is the subgroup

$$\{(\rho,\sigma) \in \mathfrak{S}(C) \times \mathfrak{S}_n \mid \rho\,\mathbf{c}\,\sigma^{-1} = \mathbf{c}\} = \{(\rho,\sigma) \in \mathfrak{S}(C) \times \mathfrak{S}_n \mid \sigma = \mathbf{c}^{-1}\rho\,\mathbf{c}\}$$

which is isomorphic to $\mathfrak{S}(C)$. It is the graph of the conjugation $\rho \mapsto \mathbf{c}^{-1}\rho\,\mathbf{c}$ which is a group isomorphism $\mathfrak{S}(C) \to \mathfrak{S}_n$.

(4) (Conjugation operation on $\mathfrak{P}(G)$) Let G be a group. Then the conjugation operation of G on G, cf. Example 1.5 (2), induces an operation of G on the power set $\mathfrak{P}(G)$ of G. For a subset A of G, the isotropy group $G_A = \{g \in G \mid gAg^{-1} = A\}$ is called the normalizer of A in G and is usually denoted by $\mathrm{N}_G(A)$. For $x \in G$, $\mathrm{N}_G(x) := \mathrm{N}_G(\{x\}) = \mathrm{C}_G(x)$ is (see Example 1.5 (2)) the centralizer of x in G. The subgroup $\mathrm{N}_G(A)$ is the biggest subgroup of G which operates on A by conjugation. The kernel of this operation of $\mathrm{N}_G(A)$ on A is the so-called centralizer $\mathrm{C}_G(A) = \bigcap_{a \in A} \mathrm{C}_G(a)$ of A. *In particular,* $\mathrm{C}_G(A)$ *is a normal subgroup of* $\mathrm{N}_G(A)$. The index of $\mathrm{N}_G(A)$ in G is the cardinality of the orbit of A which is the set of subsets of G which are conjugate to A. A subset A of G is called normal if $\mathrm{N}_G(A) = G$, or equivalently, if A is invariant under all conjugations of G. For subgroups this definition is the usual definition of normality. Since conjugates of a subgroup in G are again subgroups in G, we have: *The normalizer* $\mathrm{N}_G(H)$ *of a subgroup H in a group G is the biggest subgroup in G which contain H such that H is normal in* $\mathrm{N}_G(H)$. *The index* $[G : \mathrm{N}_G(H)]$ *is the number of conjugate subgroups of H in G and it divides the index* $[G : H]$ *of H in G if it is finite.*

(5) (Sylow's Theorems) As an application of the concept of an induced operation we prove the following theorems which were proved for the first time by P. L. M. Sylow in 1872. The following proof has been suggested by H. Wielandt.

1.6.1 Theorem (Theorems of Sylow) *Let p be a prime number and let G be a finite group of order $n = p^\alpha q$ with* $\mathrm{GCD}(p,q) = 1$. *Further, let $\beta \in \mathbb{N}$ with $\beta \leq \alpha$. Then*:

(i) *The number of subgroups of order p^β in G is congruent to* 1 *modulo p. In particular, there are subgroups in G of order p^β.*

(ii) *Let H be a subgroup of order p^α in G and let H' be a subgroup of order p^β in G. Then there exists an element $g \in G$ such that $H' \subseteq gHg^{-1}$. In particular, all subgroups of order p^α in G are conjugate.*

(iii) *The number of subgroups of order p^α in G divides q.*

For the proof consider the operation of G on the set $\mathfrak{P}_{p^\beta}(G)$ of all subsets of G of cardinality p^β induced by the left regular operation of G on G (see Example 1.5 (1) and 1.6 (3) above). Put $\gamma := \alpha - \beta$. For an $m \in \mathbb{N}^*$ we denote by $v_p(m)$ the exponent of the highest p-power which divides m. Then, for the orbit GX of an $X \in \mathfrak{P}_{p^\beta}(G)$, the following statements are equivalent:

a) $v_p(\#GX) \leq \gamma$. b) $\#GX = p^\gamma q$. c) GX contains exactly one subgroup (of order p^β).

To prove these equivalences, we consider the isotropy group G_X of X. Then $\#G_X \cdot \#GX = \#G = p^\alpha q$ and hence a) implies that $v_p(\#G_X) \geq \beta$, i. e. p^β divides $\#G_X$. For an $x \in X$, the right coset $G_X x$ is contained in X and hence $\#G_X = \#G_X x \leq \#X = p^\beta$. Altogether, we get $\#G_X = p^\beta$ and $\#GX = p^\gamma q$. For the proof of implication b) $\Rightarrow$ c) note that $\#G_X = p^\beta$ by b) and $G_X x = X$ for every $x \in X$. Therefore $x^{-1}G_X x = x^{-1}X$ is a subgroup contained in GX. Since the orbit of a subgroup H is precisely the set $G/H = \{xH \mid x \in G\}$, any orbit GX can contain at most one subgroup. This proves c). The implication c) $\Rightarrow$ a) follows from the fact that the orbit of a subgroup H of order p^β consists of the $p^\gamma q$ left cosets $G/H = \{xH \mid x \in G\}$ of H. Let d be the number of subgroups of order p^β in G. Then, by the equivalence of c) and b), d is the number of orbits of G in $\mathfrak{P}_{p^\beta}(G)$ of cardinality $p^\gamma q$. Further, by the equivalence of a) and b), the number of elements in each of the remaining orbits is divisible by $p^{\gamma+1}$. Therefore, from the class equation we get

$$\binom{p^\alpha q}{p^\beta} = \#\mathfrak{P}_{p^\beta}(G) = d \cdot p^\gamma q + r p^{\gamma+1} \qquad \text{with} \quad r \in \mathbb{N}.$$

We now apply this result to the special case of a cyclic group of order $p^\alpha q$, in which there is exactly one subgroup of order p^β (i. e. $d = 1$) and hence

$$\binom{p^\alpha q}{p^\beta} = p^\gamma q + s p^{\gamma+1} \qquad \text{with} \quad s \in \mathbb{N}.$$

Both the above equations together (by canceling p^γ) imply that

$$dq = q + (s - r)p.$$

But $\mathrm{GCD}(p,q) = 1$ and hence $d \equiv 1 \bmod p$. This proves (i). For a proof of (ii) consider the left translation operation of G on the set G/H of the set of left-cosets of H in G. We restrict this operation to the subgroup H'. Since $q = \#G/H \equiv \#\,\mathrm{Fix}_{H'}\, G/H \bmod p$ by the class equation, $\#\,\mathrm{Fix}_{H'}\, G/H \neq 0$. Therefore there exists a left coset xH which is invariant under the left translations by the elements of H', i. e. $H' \subseteq xHx^{-1}$. For a proof of (iii), by (i) there is a subgroup H of order p^α in G and by (ii) all subgroups of the order p^α are conjugates of H in G. Now, the assertion (iii) follows immediately from the last assertion in Example (4) above. □

Note that in the proof of (i) above, we proved the following result on binomial coefficients: *Let p be a prime number and let $q \in \mathbb{N}^*$ be not divisible by p, then for arbitrary $\beta, \gamma \in \mathbb{N}$, we have*:

$$\binom{p^{\beta+\gamma} q}{p^\beta} \equiv p^\gamma q \bmod p^{\gamma+1}.$$

We further remark that (i) contains as a special case ($\beta = 1$) the Theorem of Cauchy which is also proved in Example 1.5 (3).

1.7 Example (The cycle decomposition of a permutation)) Let X be a finite set and let $\sigma \in \mathfrak{S}(X)$ be a permutation of X. To describe σ perspicuously one uses the canonical operation on X of the cyclic group $\mathrm{H}(\sigma) \subseteq \mathfrak{S}(X)$ generated by σ. For an $x \in X$, the orbit $\mathrm{H}(\sigma)x$, which we call the orbit of x under σ, is the set $\{x = x_0 = \sigma^0 x, x_1 = \sigma x, \ldots, x_{m-1} =$

$\sigma^{m-1}x\}$, where m is the index of the stabilizer $\mathrm{H}(\sigma)_x$ of x in $\mathrm{H}(\sigma)$, i. e. the smallest $m \in \mathbb{N}^*$ with $x_m = \sigma^m x = x = x_0$. Therefore the operation of σ on this orbit may be written as the cycle $\langle x_0, x_1, \ldots, x_{m-1}\rangle$, where $x_0, x_1, \ldots, x_{m-1}$ are pairwise distinct elements of X with $x_{i+1} = \sigma x_i$ for $i = 0, \ldots, m-2$ and $\sigma x_{m-1} = x_0$. It is always considered as an element of $\mathfrak{S}(X)$ and has order m. If we choose a system of representatives $x^{(1)}, \ldots, x^{(r)}$ of the orbits of $\mathrm{H}(\sigma)$, then σ is the composition $\sigma_1 \cdots \sigma_r$ of the cycles $\sigma_\rho = \langle x^{(\rho)}, \sigma x^{(\rho)}, \ldots, \sigma^{m_\rho - 1} x^{(\rho)}\rangle$, where m_ρ is the index $[\mathrm{H}(\sigma) : \mathrm{H}(\sigma)_{x^{(\rho)}}]$, $\rho = 1, \ldots, r$. Observe that these cycles commute pairwise. The cycle of each fixed point is the identity and can be ignored. If X is totally ordered (e.g. $X = \{1, \ldots, n\}$), then this cycle decomposition can be normalized in the following way: $x^{(1)}$ is the smallest element in X, $x^{(2)}$ is the smallest element in $X \setminus \mathrm{H}(\sigma)x^{(1)}$ and so on.

The type $\nu(\sigma) = (\nu_i(\sigma))_{i \in \mathbb{N}^*}$ of σ characterizes the partition of $\#X$ given by the cardinalities of the orbits of σ: For an $i \in \mathbb{N}^*$, it assigns the number $\nu_i = \nu_i(\sigma)$ of orbits of cardinality i. Of course $\nu_i = 0$ for $i > \#X$. The partition itself is written as $1^{\nu_1} 2^{\nu_2} \ldots n^{\nu_n} := (1, \ldots, 1, 2, \ldots, 2, \ldots, n, \ldots, n)$, where each $i = 1, \ldots, n$ occurs ν_i times. The types of two permutations σ and τ coincide if and only if σ and τ are conjugates in $\mathfrak{S}(X)$. This is a consequence of the following general result:

1.8 Proposition *Let $\varphi : X \to Y$ be a bijective map of sets and let $\sigma \in \mathfrak{S}(X)$. Then the orbits of the "conjugate" permutation $\varphi\sigma\varphi^{-1} \in \mathfrak{S}(Y)$ are the sets $Y_\rho := \varphi(X_\rho)$, where X_ρ, $\rho = 1, \ldots, r$, are the orbits of σ. More precisely, $\varphi\langle x_0, \ldots, x_{m-1}\rangle\varphi^{-1} = \langle \varphi(x_0), \ldots, \varphi(x_{m-1})\rangle$. In particular, σ and $\varphi\sigma\varphi^{-1}$ are of the same type.*

The permutations of a given type $(\nu_i)_{i \in \mathbb{N}^*}$ in the group $\mathfrak{S}(X)$ with $\#X = n \in \mathbb{N}$ form a conjugacy class of the group $\mathfrak{S}(X)$, i. e. an orbit of the conjugation operation K: $(\tau, \sigma) \mapsto \tau\sigma\tau^{-1}$ of $\mathfrak{S}(X)$ on $\mathfrak{S}(X)$. Let $\sigma \in \mathfrak{S}(X)$ be a permutation of type $(\nu_i)_{i \in \mathbb{N}^*}$ with cycle decomposition $\sigma_1 \cdots \sigma_r$ and, for $i \in \mathbb{N}^*$, let $\mathfrak{X}_i$ be the set of orbits of σ of cardinality i. Then each $\tau \in \mathfrak{S}(X)$ with $\sigma = \tau\sigma\tau^{-1}$ induces a permutation of each $\mathfrak{X}_i$ which defines an obviously surjective group homomorphism $\mathfrak{S}(X)_\sigma \to \prod_{i=1}^n \mathfrak{S}(\mathfrak{X}_i)$. The reader easily checks that its kernel is the subgroup $\mathrm{H}(\sigma_1) \times \cdots \times \mathrm{H}(\sigma_r) \subseteq \mathfrak{S}(X)_\sigma$ generated by the cycles $\sigma_1, \ldots, \sigma_r$. Therefore $\#\,\mathfrak{S}(X)_\sigma = \prod_{i=1}^n i^{\nu_i} \cdot \prod_{i=1}^n \nu_i!$ and, by 1.3, the number of permutations of X of type $(\nu_i)_{i \in \mathbb{N}^*}$ is

$$[\mathfrak{S}(X) : \mathfrak{S}(X)_\sigma] = \frac{\#\mathfrak{S}(X)}{\#\mathfrak{S}(X)_\sigma} = \frac{n!}{1^{\nu_1}\nu_1!\,2^{\nu_2}\nu_2! \cdots n^{\nu_n}\nu_n!}\,.$$

Since $\sigma^m = \sigma_1^m \cdots \sigma_r^m$, the order[6] of σ is the LCM of the orders of the cycles $\sigma_1, \ldots, \sigma_r$:

$$\begin{aligned} \operatorname{ord}\sigma = \mathrm{LCM}(\operatorname{ord}\sigma_1, \ldots, \operatorname{ord}\sigma_r) &= \mathrm{LCM}(i \mid \nu_i(\sigma) > 0) \\ &= \mathrm{LCM}(i \mid i \geq 2,\ \nu_i(\sigma) > 0) = \mathrm{LCM}(1^{\nu_1} 2^{\nu_2} \ldots n^{\nu_n})\,. \end{aligned}$$

Following an idea of A. Cauchy, the cycle decomposition yields a convenient method to compute the sign $\operatorname{sign}\sigma$ of a permutation $\sigma \in \mathfrak{S}(X)$. Since a cycle $\langle x_0, x_1, \ldots, x_{m-1}\rangle$ of length m is the product $\langle x_0, x_1\rangle\langle x_1, x_2\rangle \cdots \langle x_{m-2}, x_{m-1}\rangle$ of $m-1$ transpositions, its sign is $(-1)^{m-1}$, hence a permutation σ of type $(\nu_i)_{i \in \mathbb{N}^*}$ has sign

$$\operatorname{sign}\sigma = (-1)^{\sum_{i \in \mathbb{N}^*} \nu_i(i-1)} = (-1)^{n-r} = (-1)^s\,,$$

where $r = \sum \nu_i$ is the number of all orbits of σ and $s := \sum \nu_{2i}$ the number of orbits of even cardinality. The sign function $\operatorname{sign} : \mathfrak{S}(X) \to \{\pm 1\}$, $\sigma \mapsto \operatorname{sign}\sigma$, is a group homomorphism. Its kernel is the alternating group $\mathfrak{A}(X)$, which is of index 2 in $\mathfrak{S}(X)$ if $\#X > 1$. By definition, the elements of $\mathfrak{A}(X)$ are called the even permutations of X.

---

[6] Let us recall that the order $\operatorname{ord} a$ of an element a in a group G is the unique non-negative generator of the kernel of the exponential homomorphism $\mathbb{Z} \to G$, $n \mapsto a^n$. This order is 0 if and only if this group homomorphism is injective, in this case many authors say that a is of infinite order. Otherwise $\operatorname{ord} a$ is the order of the subgroup $\mathrm{H}(a)$ of G generated by a, i. e. the image of the exponential map.

1.9 Example (Simply transitive operations) The most natural simply transitive operation of a group G is the operation of G on G by left translations. The operation map $G \times G \to G$ is the multiplication in the group G and the action homomorphism $G \to \mathfrak{S}(G)$ maps g to the left multiplication $\mathrm{L}_g : G \to G, x \mapsto gx$. This operation is called the regular or Cayley operation of G, cf. Example 1.5 (1). The action homomorphism identifies G with a subgroup of the permutation group $\mathfrak{S}(G)$ and realizes G as a transformation group, see Footnote 5. Up to isomorphism the Cayley operation of G is the only simply transitive operation of G: *If $G \times X \to X$ is simply transitive and if $x_0 \in X$ is an arbitrary point in X, then $G \to X$, $g \mapsto gx_0$, is a bijective G-morphism if G operates on G regularly.*

The most classical examples of simply transitive group operations are given by the affine spaces. By definition an affine space is a set E with a simply transitive operation of the additive group of a vector space V (over a field K). Usually this operation is also written additively: $(v,P) \mapsto v+P$, $v \in V$, $P \in E$. This definition of affine spaces goes back to Hermann Weyl, see his famous book "Space, Time, Matter". The unique vector $v \in V$ which transforms the point $P \in E$ to the point $Q \in E$ is denoted by $\overrightarrow{PQ}$, i. e. $\overrightarrow{PQ}+P = Q$. Since the group V is commutative we can change to the more appealing notation $P+\overrightarrow{PQ} = Q$.

With these examples in mind, any G-space X with a simply transitive operation of a group G is called a G-affine space. For $x,y \in X$, the unique element $g \in G$ with $y = gx$ is denoted by $\overrightarrow{xy}$ and G is called the group of translations. As mentioned above a G-affine space X can be identified with the group G of translations *if an origin $x_0 \in X$ is chosen.* (This choice is arbitrary but not canonical. A group has an exceptional element, namely its identity element e_G, but an affine space does not have such an element.) The set $\mathfrak{P}_C$ of packs of 1.1 is an $\mathfrak{S}_n$-affine space, $n := \#C$.

The right multiplications $\mathrm{R}_g : G \to G, x \mapsto xg$, define the regular right operation of G on G. To transform it into a left operation one has to use the action homomorphism $g \mapsto \mathrm{R}_{g^{-1}}$ from $G \to \mathfrak{S}(G)$. Put together, both regular operations give the (left) operation of the product group $G \times G$ on G by $((g,h),x) \mapsto gxh^{-1}$. At the end of Example 1.6(3), if we identify $\mathfrak{S}(C)$ with $\mathfrak{S}_n$ (which is not canonical and we do not like to do this), then the operation there is isomorphic to the operation of $\mathfrak{S}_n \times \mathfrak{S}_n$ on $\mathfrak{S}_n$ considered here.

1.10 Example (Transitive operations) For every subgroup $H \subseteq G$, the (left) Cayley operation of a group G on itself induces a transitive operation on the set G/H of left cosets of H in G (which, by the way, is the set of orbits of the canonical operation of H on G from the right), i. e. G/H is the orbit of H of the operation of G on $\mathfrak{P}(G)$ induced by the Cayley operation, see Example 1.6 (3). The stabilizer of $H = e_G H$ is the group H itself and hence the stabilizer of $gH \in G/H$ is the conjugate group $gHg^{-1}, g \in G$, see 1.3. From the view point of group operations the special coset H cannot be distinguished from any other left coset gH, i. e. only the conjugacy class of H is an invariant as the set of isotropy groups of this operation.

More generally, let X be a homogeneous G-space and let $x_0 \in X$ be an arbitrary chosen point of X (as an origin) with stabilizer $H \subseteq G$. Then $gH \mapsto gx_0$ is a well-defined bijective map from G/H to X. Indeed, it is even a G-isomorphism, where G/H carries the canonical G-operation mentioned above. If we make another choice $y_0 = g_0 x_0$ for the origin, then we have to replace H by the conjugate $g_0 H g_0^{-1}$. We call the conjugacy class of H also the stabilizer of the homogeneous G-space X. The kernel of the operation of G on G/H is the intersection $\mathrm{K}_G(H) := \bigcap_{g \in G} gHg^{-1}$ of the conjugates of H. It is the biggest normal subgroup of G contained in H. The operation of G induces a faithful operation of $G/\mathrm{K}_G(H)$ on G/H. In particular, the operation of G on G/H is faithful if and only if $\mathrm{K}_G(H) = \{e_G\}$ and, in general, $G/\mathrm{K}_G(H)$ is isomorphic to a subgroup of $\mathfrak{S}(G/H)$. Especially, if H is of finite index in G, then $\mathrm{K}_G(H)$ is of finite index too. If $G \neq \{e_G\}$

itself is finite and if, moreover, $\#(G/H) = p$ is the smallest prime divisor of $\#G$, then $H = \mathrm{K}_G(H)$ is necessarily normal.

We summarize these considerations in the following theorem:

1.11 Theorem *Up to isomorphism the homogeneous G-spaces are given by the G-spaces G/H, H subgroup of G, with the canonical G-operations. The kernel of the operation of G on G/H is $\mathrm{K}_G(H) = \bigcap_{g\in G} gHg^{-1}$. – Two such homogeneous spaces $G/H\,, G/H'$ are G-isomorphic if and only if H and H' are conjugate subgroups in G. More generally, for a given $\varphi \in \mathrm{Aut}\, G$, there exists a φ-isomorphism $G/H \to G/H'$ if and only if $\varphi(H)$ and H' are conjugate subgroups in G.*

The conjugation in G is an operation of the group G on the set G, indeed it is the restriction of the Cayley operation of the product group $G \times G$ on G by left and right translations (see the end of Example 1.9) to the diagonal subgroup $\Delta_G = \{(g,g) \mid g \in G\} \subseteq G \times G$. It induces the conjugation on the set of subgroups of G. Theorem 1.11 says that the orbits of the last operations are in one-to-one correspondence to the isomorphism classes of homogeneous G-spaces. Therefore, *an arbitrary G-space is up to isomorphism characterized by its* type or Burnside function. This function assigns to each conjugacy class of subgroups of G the number of those orbits which have the given class as their stabilizer. See also the article On the Burnside Algebra of a Finite Group, Dilip P Patil and Anshoo Tandon, pp 103-120, in this Journal.

To determine the G-automorphisms of a homogeneous G-space G/H, let $f : G/H \to G/H$ be such a G-automorphism and let $f(H) = g_0H$ for some $g_0 \in G$. Then $f(gH) = gf(H) = gg_0H$ for all $g \in G$. In particular, $g_0H = f(H) = f(hH) = hf(H) = hg_0H$ for all $h \in H$ and hence $H \subseteq g_0Hg_0^{-1}$, since $g_0Hg_0^{-1}$ is the stabilizer of $g_0H \in G/H$. The inverse $f^{-1} : G/H \to G/H$ maps g_0H to H and hence $H = g_0^{-1}(g_0H)$ to $g_0^{-1}H$ which implies $H \subseteq g_0^{-1}Hg_0$. Altogether we have the equality $H = g_0Hg_0^{-1}$. Conversely, if $g_0 \in G$ is an element with $H = g_0Hg_0^{-1}$, then, obviously, the map $f_{g_0} : gH \mapsto gg_0H$ is well-defined and a G-automorphism of G/H. The map $g_0 \mapsto f_{g_0}$ is a surjective anti-homomorphism from $\mathrm{N}_G(H) := \{g_0 \in G \mid g_0Hg_0^{-1} = H\}$ to the group $\mathrm{Aut}_G(G/H)$ of G-automorphisms of G/H. The kernel of this anti-homomorphism is H and hence the groups $\mathrm{Aut}_G(G/H)$ and $\mathrm{N}_G(H)/H$ are anti-isomorphic (and hence isomorphic). The subgroup $\mathrm{N}_G(H) \subseteq G$ is called the normalizer of H in G. It is the biggest subgroup of G such that H is normal in $\mathrm{N}_G(H)$ and occured already in Example 1.6 (4). If H is already normal in G, then $\mathrm{N}_G(H) = G$ and $\mathrm{Aut}_G(G/H) \cong (G/H)^{\mathrm{op}}$ $(\cong G/H)$. In particular, if $H = \{e_G\}$, i. e. if the operation is simply transitive, then $\mathrm{Aut}_G(G) \cong G^{\mathrm{op}}$. Particularly, the G-automorphisms of G are the right translations R_{g_0} of G, $g_0 \in G$.

1.12 Example (Affine Groups) An affine automorphism of an affine space E over a K-vector space V is, by definition, a φ-automorphism for some $\varphi \in \mathrm{GL}_K(V)$. The group $\mathrm{Aff}_K(E)$ of these automorphisms is called the affine group of the affine space E. The V-automorphisms, i. e. the id_V-automorphisms, are the translations ϑ_v of E, $v \in V$. They form a normal subgroup $\mathrm{T}(E)$ of $\mathrm{Aff}_K(E)$, which is isomorphic to the additive group of V. The map $\mathrm{Aff}_K(E) \to \mathrm{GL}_K(V)$ which maps a φ-automorphism of E to the K-linear automorphism φ of V is surjective with kernel $\mathrm{T}(E)$, i. e. $\mathrm{Aff}_K(E)/\mathrm{T}(E) \cong \mathrm{GL}_K(V)$. Every affine automorphism of E is a collineation of E, where, by definition, a collineation of the affine space E is a bijective map $E \to E$ which maps affine lines in E onto affine lines in E. Since, obviously, the inverse of a collineation is also a collineation, the set $\mathrm{Coll}_K(E)$ of all collineations of E is a subgroup of the permutation group $\mathfrak{S}(E)$. For example, $\mathrm{Coll}_K(E) = \mathfrak{S}(E)$ if $\mathrm{Dim}_K E = 1$. By the following theorem collineations and affine automorphisms coincide in important cases:

1.12.1 Theorem (Fundamental Theorem of Affine Geometry) *If the automorphism group* $\mathrm{Aut}\, K$ *of the field K is trivial,* i. e. *the identity* id_K *is the only automorphism of the*

field K, and if $\mathrm{Dim}_K E = \mathrm{Dim}_K V \geq 2$*, then every collineation of the affine space E over K is an affine automorphism,* i. e. $\mathrm{Coll}_K(E) = \mathrm{Aff}_K(E)$.

For example, for a prime field $\mathbb{Q}$, respectively $\mathbb{F}_p (\cong \mathbb{Z}_p = \mathbb{Z}/\mathbb{Z}p)$, p a prime number, as well as for the field $\mathbb{R}$ of real numbers, the automorphism group is trivial and hence the above theorem holds. For $K = \mathbb{R}$ and $\mathrm{Dim}_{\mathbb{R}} E = 3$, this was already known to Euler. – More generally, if $\mathrm{Dim}_K E \geq 2$, then the group $\mathrm{Aff}_K(E)$ is a normal subgroup of the group $\mathrm{Coll}_K(E)$ with residue class group $\mathrm{Coll}_K(E)/\mathrm{Aff}_K(E) \cong \mathrm{Aut}\, K$. For a proof see for instance [5, Teil 1, § 43, Theorem 43.8].

In the spirit of the last discussion, the affine automorphisms of an arbitrary G-affine space X are, by definition, the φ-automorphisms of X, where φ runs through the full group $\mathrm{Aut}\, G$ of group automorphisms of G. The φ-automorphism $f : X \to X$ has by definition the property $f(x) = f(\overrightarrow{x_0x} \cdot x_0) = \varphi(\overrightarrow{x_0x}) \cdot f(x_0)$ or $\overrightarrow{f(x_0)f(x)} = \varphi(\overrightarrow{x_0x})$ for arbitrary $x, x_0 \in X$. Conversely, given an arbitrary point $y_0 (= f(x_0)) \in X$, the map $x \mapsto \varphi(\overrightarrow{x_0x})y_0$ is a φ-automorphism f of X with $x_0 \mapsto y_0$ which is uniquely determined by φ and the image $y_0 = f(x_0)$ of a single point $x_0 \in X$.

The surjective map $\mathrm{Aff}(X) \to \mathrm{Aut}\, G$ of the group $\mathrm{Aff}(X)$ of affine automorphisms of X onto the automorphism group $\mathrm{Aut}\, G$ which maps an affine automorphism $f \in \mathrm{Aff}(X)$ to the group automorphism $\varphi \in \mathrm{Aut}\, G$ as described above is a group homomorphism with the group $\mathrm{T}(X)$ of G-automorphisms of X (which are also called the translations of X) as kernel. The group $\mathrm{T}(X)$ is isomorphic to $G^{\mathrm{op}} (\cong G)$, cf. the end of Example 1.10.

For a fixed point $x_0 \in X$, chosen as origin, the subgroup $\mathrm{Aff}_{x_0}(X) \subseteq \mathrm{Aff}(X)$ of affine automorphisms with fixed point x_0 maps isomorphically to $\mathrm{Aut}\, G$. In particular, the exact sequence

$$1 \to \mathrm{T}(X) \to \mathrm{Aff}(X) \to \mathrm{Aut}\, G \to 1$$

of groups splits (weakly), i. e. $\mathrm{Aff}(X)$ is a semidirect product $\mathrm{T}(X) \rtimes \mathrm{Aut}\, G$. The operation of $\mathrm{Aut}(G) = \mathrm{Aut}(G^{\mathrm{op}})$ on $\mathrm{T}(X)$ in this semidirect product is the natural one given by the identification $\mathrm{T}(X) = G^{\mathrm{op}}$ (depending on the chosen origin x_0): For $\varphi \in \mathrm{Aut}\, G$, the corresponding element $f_0 \in \mathrm{Aff}_{x_0}(X)$ is $x \mapsto \varphi(\overrightarrow{x_0x})x_0$. The translation T_g corresponding to $g \in G^{\mathrm{op}}$ is $x \mapsto (\overrightarrow{x_0x})gx_0$, and hence $f_0 \mathrm{T}_g f_0^{-1} = \mathrm{T}_{\varphi(g)}$, since $(f_0\mathrm{T}_g)(x_0) = f_0(gx_0) = \varphi(g)f(x_0) = (\mathrm{T}_{\varphi(g)}f_0)(x_0)$.

An (abstract) semidirect product is constructed from an operation of a group H on a group G *by group automorphisms*, i. e. the image of the action homomorphism $\vartheta : H \to \mathfrak{S}(G)$ is contained in $\mathrm{Aut}\, G$. Then $G \times H$ with the binary operation

$$(g,h) \cdot (g',h') := (g(\vartheta_h g'), hh')$$

is a group $G \rtimes H = G \rtimes_\vartheta H$ and $\{(g, e_H) \mid g \in G\}$ is a normal subgroup identified with G and $\{(e_G, h) \mid h \in H\}$ is a subgroup identified with H. The projection map $(g,h) \mapsto h$ is a surjective group homomorphism $G \rtimes_\vartheta H \to H$ with kernel G. Hence, there is a canonical exact sequence

$$1 \to G \to G \rtimes_\vartheta H \to H \to 1$$

which splits (weakly) and the conjugation $hgh^{-1} = \vartheta_h(g)$, $g \in G, h \in H$, of H on G in the semidirect product $G \rtimes_\vartheta H$ is the given operation ϑ.

In the special case where X is the canonical G-affine space G with the Cayley operation, the affine group $\mathrm{Aff}(G) \subseteq \mathfrak{S}(G)$ is called the (full) holomorph $\mathrm{Hol}(G)$ of G. It is generated by the right translations $\mathrm{R}_g, g \in G$, and the automorphisms of G. Because of $\mathrm{L}_g \circ \mathrm{R}_g^{-1} = \mathrm{K}_g \in \mathrm{Aut}\, G$, where $\mathrm{K}_g : G \to G$ is the conjugation with g, $\mathrm{Hol}(G)$ contains also the left translations $\mathrm{L}_g, g \in G$, of G and $\mathrm{Hol}(G) = \mathrm{Hol}(G^{\mathrm{op}})$. $\mathrm{Hol}(G)$ is the semidirect product $G \rtimes \mathrm{Aut}\, G$ with

respect to the canonical operation of Aut G on G. The map $(g,\varphi)\mapsto \mathrm{L}_g\circ\varphi$ is an isomorphism $G\rtimes \mathrm{Aut}\,G\to \mathrm{Hol}(G)$. If one allows for the automorphisms $\varphi\in \mathrm{Aut}\,G$ only elements of a given subgroup $\Phi\subseteq \mathrm{Aut}\,G$, one obtains a subgroup $\mathrm{Hol}_\Phi(G)$ of the full holomorph $\mathrm{Hol}(G)$ which is canonically isomorphic to $G\rtimes\Phi$. This is already done for the affine spaces over K-vector spaces V, where φ is restricted to the K-linear automorphisms of V. Therefore $\mathrm{Aff}_K(V)=\mathrm{Hol}_{\mathrm{GL}_K(V)}(V)$ is a subgroup of the full holomorph $\mathrm{Hol}(V)$ of the additive group V. If K is a prime field then $\mathrm{Aff}_K(V)=\mathrm{Hol}(V)$.

1.13 Example Very often the determination of the orbit space $X\backslash G$ of a G-space X can be considered as a classification problem: Two objects $x,y\in X$ are considered as equivalent (or isomorphic or indistinguishable or of the same structure or ...) if $y=gx$ for some $g\in G$, i. e. if they belong to the same orbit. We consider two examples.

First, consider the (K-linear) endomorphisms f of a K-vector space V of finite dimension n. If one chooses a basis $\mathfrak{v}=(v_1,\dots,v_n)$ of V, then f is uniquely determined by its matrix $\mathfrak{A}=(a_{ij})_{1\le i,j\le n}$ with respect to the basis $\mathfrak{v}$ which is defined by the equations $f(v_j)=\sum_{i=1}^n a_{ij}v_i$ for $j=1,\dots,n$. Change of the basis to $\mathfrak{w}=(w_1,\dots,w_n)$ transfers the matrix into the conjugated matrix $\mathfrak{B}=\mathfrak{G}\mathfrak{A}\mathfrak{G}^{-1}$, where $\mathfrak{G}=(g_{ij})\in \mathrm{GL}_n(K)$ describes the base change: $v_j=\sum_{i=1}^n g_{ij}w_i$, $j=1,\dots,n$. Therefore, to classify the endomorphisms of V means to describe in a simple way the orbits of $\mathrm{M}_n(K)$ under the conjugation operation of the group $\mathrm{GL}_n(K)$. In case K is algebraically closed this is done by the well-known Jordan block matrices (which can be generalized suitably to arbitrary fields).

Secondly, let M be a nice topological space, for instance, a connected topological manifold. The connected covering spaces $p:\widetilde{M}\to M$ correspond up to isomorphism to the homogeneous $\pi(M)^{\mathrm{op}}$-spaces, where $\pi(M)=\pi(M,P)$ is the fundamental group of M with respect to an arbitrary base point $P\in M$: The canonical operation of the group $\pi(M)$ from the *right* on the fibre $p^{-1}(P)\subseteq\widetilde{M}$ over P is transitive (since $\widetilde{M}$ is connected) and characterizes $\widetilde{M}$ up to isomorphism. Therefore, by Theorem 1.11, the connected covering spaces of M are classified by the orbits of the subgroups of $\pi(M)^{\mathrm{op}}$ or $\pi(M)$ under conjugation. If a given covering $p:\widetilde{M}\to M$ corresponds to the conjugacy class of $H^{\mathrm{op}}\subseteq\pi(M)^{\mathrm{op}}$, then the automorphism group $\mathrm{Deck}(\widetilde{M},p)=\{F:\widetilde{M}\to\widetilde{M}\mid F \text{ continuouswith } p\circ F=p\}$ (which is called the deck transformation group of the covering p) can be identified with the $\pi(M)^{\mathrm{op}}$-automorphism group of the homogeneous space $\pi(M)^{\mathrm{op}}/H^{\mathrm{op}}$, which is, by the end of Example 1.10, isomorphic to $\mathrm{N}_{\pi(M)^{\mathrm{op}}}(H^{\mathrm{op}})/H^{\mathrm{op}}=(\mathrm{N}_{\pi(M)}(H)/H)^{\mathrm{op}}\cong \mathrm{N}_{\pi(M)}(H)/H$. If H is a normal subgroup of $\pi(M)$, then this group is $(\pi(M)/H)^{\mathrm{op}}\cong\pi(M)/H$ (and the deck transformation group operates transitively on the fibres of the covering).

1.14 Example (Galois operations) Let K be a field and let $F\in K[X]$ be a monic irreducible and separable polynomial of degree n with coefficients in K. By a classical result of Kronecker, there exists a finite field extension L of K such that the polynomial F splits into linear factors over L, i. e. in $L[X]$ one has $F=(X-\alpha_1)\cdots(X-\alpha_n)$ with pairwise distinct $\alpha_1,\dots,\alpha_n\in L$ (since F is separable). We may assume that $L=K[\alpha_1,\dots,\alpha_n]$ is generated over K by the zeros $\alpha_1,\dots,\alpha_n$ of F. Then L is called the (minimal)splitting field of F over K. It is a so-called Galois extension and uniquely determined by F up to K-algebra isomorphism. The group $\mathrm{G}(L|K)$ of K-algebra automorphisms of L is called the Galois group $\mathrm{Gal}_K(F)$ of the polynomial F over K. Since the image $\varphi(\alpha)$ of any zero α of F under a K-algebra homomorphism φ is again a zero of F, the canonical operation of $\mathrm{Gal}_K(F)$ on L induces an operation of $\mathrm{Gal}_K(F)$ on the finite set $\mathrm{V}(F)=\{\alpha_1,\dots,\alpha_n\}$. Since any K-algebra homomorphism of L is uniquely determined by the images of the K-generators $\alpha_1,\dots,\alpha_n$, this induced operation of $\mathrm{Gal}_K(F)$ is also faithful. Furthermore, it is transitive which is a consequence of the irreducibility of F. Therefore the Galois group of F is a group G which operates transitively and faithfully

on a finite set X of n elements, i. e. $\mathrm{Gal}_K(F)$ *is isomorphic to a subgroup of the permutation group* $\mathfrak{S}_n$ *for which the canonical operation on* $\{1,\dots,n\}$ *is transitive.* This identification of $\mathrm{Gal}_K(F)$ as a subgroup of $\mathfrak{S}_n$ determines $\mathrm{Gal}_K(F)$ only up to conjugation in $\mathfrak{S}_n$. Therefore, *the classification of the Galois operations defined by irreducible separable polynomials of degree n is equivalent to the classification of the conjugacy classes of the transitive subgroups of* $\mathfrak{S}_n$.

By Galois theory, the order of $\mathrm{Gal}_K(F) = \mathrm{G}(L|K)$ is the degree $[L:K] = \mathrm{Dim}_K L$ of the splitting field L of F over K. It divides $n! = (\deg F)!$ and is divisible by n. The quotient $n!/\#\mathrm{Gal}_K(F)$, i. e. the index of $\mathrm{Gal}_K(F)$ in $\mathfrak{S}(\mathrm{V}(F)) \cong \mathfrak{S}_n$, is called the affect of F. If the affect is 1, i. e. if $\mathrm{Gal}_K(F) \cong \mathfrak{S}_n$, then the polynomial F is called affectless or without affect. On the other hand, the order of $\mathrm{Gal}_K(F)$ is $n = \deg(F)$ if and only if $\mathrm{Gal}_K(F)$ operates simply transitively on $\mathrm{V}(F)$, or, equivalently, $L = K[\alpha_1] \cong K[X]/(F)$, i. e. if F splits already over $K[\alpha_1]$ into linear factors. In this case F is called a Galois polynomial. For instance, if $\mathrm{Gal}_K(F)$ is abelian, then F is necessarily a Galois polynomial. The splitting field L is always generated by one element over K, i. e. $L = K[\alpha]$ for some $\alpha \in L$. This is a special case of the primitive element theorem. Such a generator α, is called a (Galois) resolvent of the polynomial F and its minimal polynomial $R \in K[X]$ which is of degree $\#\mathrm{Gal}_K(F)$ is called the resolvent (Galois) polynomial corresponding to the chosen resolvent α. The resolvent polynomial R splits completely over the field L into linear factors. Hence R is a Galois polynomial with Galois group $\mathrm{Gal}_K(R)$, which is isomorphic to $\mathrm{Gal}_K(F)$. But the representation in $\mathfrak{S}_m$, $m := \#\mathrm{Gal}_K(F)$, is now simply transitive.

Instead of looking for representations of Galois groups in permutation groups, one can ask for a given finite group G, what the possible degrees of separable irreducible polynomials with Galois group (isomorphic to) G are. By Theorem 1.11, these are exactly the indices $[G:H]$ of those subgroups $H \subseteq G$ for which $\mathrm{K}_G(H) = \bigcap_{g\in G} gHg^{-1} = \{e_G\}$, where two such subgroups H, H' have to be identified if there is an automorphism (not necessarily an inner automorphism) $\psi : G \to G$ with $\psi(H) = H'$.

For $1 \le n \le 5$, the conjugacy classes of transitive subgroups G of $\mathfrak{S}_n$ are represented by the following groups (as one can check more or less straightforwardly, we put $\gamma_n := \langle 1,2,\dots,n\rangle$):

$n = 1$: $\{\gamma_1\} = \mathfrak{S}_1$.

$n = 2$: $\mathfrak{S}_2$.

$n = 3$: $\mathfrak{S}_3,\ \mathfrak{A}_3 = \mathrm{H}(\gamma_3)$.

$n = 4$: $\mathfrak{S}_4,\ \mathfrak{A}_4,\ \mathbf{D}_4 = \mathrm{H}(\gamma_4, \langle 2,4\rangle),\ \mathfrak{V}_4 = \{\mathrm{id}_4, \langle 1,2\rangle\langle 3,4\rangle, \langle 1,3\rangle\langle 2,4\rangle, \langle 1,4\rangle\langle 2,3\rangle\}$, $\mathbf{Z}_4 = \mathrm{H}(\gamma_4)$. (The classes of $\mathbf{D}_4$ and $\mathbf{Z}_4$ contain 3 conjugate subgroups each.)

$n = 5$: $\mathfrak{S}_5,\ \mathfrak{A}_5,\ \mathrm{Hol}(\mathbf{Z}_5) = \mathrm{H}(\gamma_5, \langle 1,2,4,3\rangle)$, $\mathbf{D}_5 = \mathrm{Hol}_{\{\pm 1\}}(\mathbf{Z}_5) = \mathrm{H}(\gamma_5, \langle 1,4\rangle\langle 2,3\rangle)$, $\mathbf{Z}_5 = \mathrm{H}(\gamma_5)$. (The classes of $\mathrm{Hol}(\mathbf{Z}_5)$, $\mathbf{D}_5$ and $\mathbf{Z}_5$ contain 6 conjugate subgroups each For the holomorph of groups see Example 1.12.)

For non-commutative groups G up to order 8, the homogeneous spaces G/H with faithful action homomorphism are represented by the following groups[7]:

$G = \mathfrak{S}_3 \cong \mathbf{D}_3 = \mathrm{Hol}(\mathbf{Z}_3) = \mathbf{Z}_3 \rtimes \{\pm 1\}$: $H = \{(0,\pm 1)\},\ H = \{(0,1)\}$,

$G = \mathbf{D}_4 = \mathrm{Hol}(\mathbf{Z}_4) = \mathbf{Z}_4 \rtimes \{\pm 1\}$: $H = \{(0,\pm 1)\},\ H = \{(0,1)\}$,

[7]Note that, for an abelian group G, the only faithful transitive operation is (up to isomorphism) the simply transitive Cayley operation of Example 1.5 (1).

(Note that all subgroups $H \subseteq \mathbf{D}_4$ of order 2 for which $\mathbf{D}_4$ acts faithfully on $\mathbf{D}_4/H$ are equivalent with respect to $\operatorname{Aut}\mathbf{D}_4$ but they form two conjugacy classes.)

$$G = \mathbf{Q}\,(= \mathbf{Q}_2 \cong \mathbb{H}(\mathbb{Z})^\times = \{\pm 1, \pm \mathrm{i}, \pm \mathrm{j}, \pm \mathrm{k}\}) \quad : \quad H = \{1\}.$$

(The group $\mathbf{Q}$ is the quaternion group of order 8. It is non-commutative, but all its subgroups are normal. All separable irreducible polynomials with Galois group $\mathbf{Q}$ are necessarily Galois polynomials (of degree 8).)

In general, for a dihedral group $\mathbf{D}_n = \operatorname{Hol}_{\{\pm 1\}}(\mathbf{Z}_n) = \mathbf{Z}_n \rtimes \{\pm 1\}$, $n \geq 3$, the subgroups $\{(0, \pm 1)\}$ and $\{(0,1)\}$ represent the $\operatorname{Aut}\mathbf{D}_n$-equivalence classes of subgroups H with a faithful canonical operation of $\mathbf{D}_n$ on $\mathbf{D}_n/H$. In particular, a separable irreducible polynomial with Galois group $\mathbf{D}_n$, $n \geq 3$, has degree n or $2n$. For $n = 1, 2$, the degree is necessarily $2n$. (Note that $\mathbf{D}_1 \cong \mathbf{Z}_2$ and $\mathbf{D}_2 \cong \mathbf{Z}_2 \times \mathbf{Z}_2$ are commutative.)

We remark that there are isomorphic transitive subgroups $G, G' \subseteq \mathfrak{S}_n$ which are not conjugates in $\mathfrak{S}_n$. The above list shows that in this case n is necessarily > 5. To construct a simple example consider the subgroups $H := \mathrm{H}(\langle 1,2\rangle)$ and $H' := \mathrm{H}(\langle 1,2\rangle\langle 3,4\rangle)$ in $\mathfrak{S}_4$ which are not conjugates. The group $\mathfrak{S}_4$ acts transitively on $\mathfrak{S}_4/H$ and $\mathfrak{S}_4/H'$. By Theorem 1.11, these two homogeneous spaces of cardinality 12 are not isomorphic. (Observe that *all* automorphisms of $\mathfrak{S}_4$ are inner!) Identifying both sets $\mathfrak{S}_4/H$ and $\mathfrak{S}_4/H'$ with $\{1, 2, \ldots, 12\}$, we get transitive and faithful actions of $\mathfrak{S}_4$ on $\{1, 2, \ldots, 12\}$. The images of the corresponding action homomorphisms are subgroups $G, G' \subseteq \mathfrak{S}_{12}$ which are both isomorphic to $\mathfrak{S}_4$, but not conjugates in $\mathfrak{S}_{12}$.

Finally, we note that any transitive subgroup $G \subseteq \mathfrak{S}_n$ can be realized as the Galois action on the set of zeros of an irreducible separable polynomial F over an appropriate field K. For this start with any polynomial P of degree n without affect over a field k (for instance, with the monic generic polynomial $P := X^n + U_1 X^{n-1} + \cdots + U_{n-1} X + U_n$ over a rational function field $k := k_0(U_1, \ldots, U_n)$, where $U_1, \ldots, U_n$ are indeterminates over an arbitrary field k_0). Let L be the splitting field of P over k and let $\alpha_1, \ldots, \alpha_n$ be the zeroes of P in L. With this enumeration of the zeros of P, the Galois group $\operatorname{Gal}_k(P) = \mathrm{G}(L|k)$ can be identified with $\mathfrak{S}_n$. Then $K := \operatorname{Fix}_G L$ and P considered as a polynomial in $K[X] \supseteq k[X]$ is a realization of $G \cong \operatorname{Gal}_K(P) = \mathrm{G}(L|K)$ as a Galois group for an irreducible separable polynomial of degree n. This is a simple consequence of the basics of Galois theory.

Usually, it is a difficult problem to decide which Galois groups can be realized over a given field K – or a bit stronger – which transitive and faithful operations of finite groups can be realized, up to isomorphism, as Galois operations on the zeros of an irreducible and separable polynomial over K (as described in the beginning of this example). Problems of this kind belong to the so-called Inverse Galois Theory. For example, it is an unsolved problem whether or not every finite group is isomorphic to the Galois group of a finite Galois extension of $\mathbb{Q}$. – For the rational function field $\mathbb{C}(t)$, t an indeterminate over the field of complex numbers $\mathbb{C}$, each finite group G occurs up to isomorphism as a Galois group of a finite extension of $\mathbb{C}(t)$ even in the strong sense that a transitive and faithful operation of G can be prescribed. This follows easily from the theory of compact Riemann surfaces by interpreting $\mathbb{C}(t)$ as the field of meromorphic functions of the Riemann sphere $\mathbb{P}^1(\mathbb{C})$ and by using the theory of covering spaces as described in the second part of Example 1.13. (See for instance [7, Bd. 4, Example 16.A.4].) – Over a finite field K a finite group G can be realized up to isomorphism as a Galois group of a finite extension L of K if and only if G is cyclic. *The Galois group* $\mathrm{G}(L|K)$ *has even a canonical generator*, namely the Frobenius automorphism $x \mapsto x^q$ of L, $q := \#K$.

§ 2 Elementary Number Theory

In this section we collect some results from elementary number theory which will be used later. The most important objects, which are studied in elementary number theory are the residue class rings $\mathbb{Z}_m := \mathbb{Z}/\mathbb{Z}m$, $m \in \mathbb{N}$. Integers $a, b \in \mathbb{Z}$ which represent the same element in $\mathbb{Z}_m$, i. e. with m divides $b-a$, are called congruent modulo m. This is denoted by $a \equiv b \bmod m$. For $m = 0$ the ring $\mathbb{Z}_m$ is $\mathbb{Z}$ itself. For $m > 0$ the residue classes $[a]_m \in \mathbb{Z}_m$, $a \in \mathbb{Z}$, have the canonical system of representatives $0, 1, \ldots, m-1$ given by the classical Euclidean division algorithm for integers. In particular, $\#\mathbb{Z}_m = m$ for $m > 0$. For an arbitrary ring A, the map $a \mapsto a \cdot 1_A$ is the only *ring* homomorphism $\chi_A : \mathbb{Z} \to A$. Its kernel is generated uniquely by the non-negative integer $m := \mathrm{Char}A := \mathrm{ord}\, 1_A$ (= the order of 1_A in the additive group of A) and its image, which is isomorphic to the quotient ring $\mathbb{Z}_{\mathrm{Char}A}$, is the smallest subring of A. Hence $\mathbb{Z}_m$, $m \in \mathbb{N}$, is up to isomorphism the smallest ring of characteristic $m \in \mathbb{N}$. Therefore it is also called the prime ring of characteristic m.[8] For every multiple $n \in \mathbb{N}$ of m, the surjective ring homomorphism $\chi : \mathbb{Z} \to \mathbb{Z}_m$ induces a surjective ring homomorphism $\overline{\chi} : \mathbb{Z}_n \to \mathbb{Z}_m$, $[a]_n \mapsto [a]_m$, with kernel $\mathbb{Z}m/\mathbb{Z}n$.

The additive group of $\mathbb{Z}_m$ is easily described. It is a cyclic group $\mathbf{Z}_m$ which is infinite for $m = 0$ and of order m for $m > 0$ and which is generated by $[1]_m$ or, more generally, by any element $[a]_m$, $a \in \mathbb{Z}$, with $\mathrm{GCD}(a, m) = 1$. These elements are precisely the units in the ring $\mathbb{Z}_m$ and are called the prime residue classes modulo m. They form the (multiplicative) group

$$\mathbb{Z}_m^\times$$

of units in $\mathbb{Z}_m$. For $m = 0$, the group $\mathbb{Z}_0^\times = \mathbb{Z}^\times = \{\pm 1\}$ is cyclic of order 2. For $m > 0$, the order of the group $\mathbb{Z}_m^\times$ is

$$\varphi(m) := \#\mathbb{Z}_m^\times = \#\{a \in \mathbb{N} \mid 0 \le a < m,\ \mathrm{GCD}(a, m) = 1\}.$$

The function $\varphi : \mathbb{N}^* \to \mathbb{N}$, $m \mapsto \varphi(m)$, is called the Euler function. From Lagrange's Theorem, it follows

$$\text{Euler's Equation}: [a]_m^{\varphi(m)} = [1]_m, \quad \text{i. e.} \quad a^{\varphi(m)} \equiv 1 \bmod m, \text{ if } \mathrm{GCD}(a, m) = 1.$$

In particular, for a prime number p,

$$a^{p-1} \equiv 1 \bmod p \quad \text{if} \quad \mathrm{GCD}(a, p) = 1,$$

which is known as Fermat's Little Theorem. For a more precise description of the unit group $\mathbb{Z}_m^\times$ the so called Chinese Remainder Theorem is useful.

2.1 Chinese Remainder Theorem *Let $m_1, \ldots, m_r$ be pairwise relatively prime positive integers and let $m := m_1 \cdots m_r$. Then for every r-tuple $(a_1, \ldots, a_r) \in \mathbb{Z}^r$ of integers, the*

[8] Note the difference between $\mathbf{Z}_m$ denoting a cyclic *group* of order m if $m > 0$ and of infinite order if $m = 0$, and $\mathbb{Z}_m$ denoting a prime *ring* of characteristic m. – For an integer $a \in \mathbb{Z}$, we also write a for the element $a \cdot 1_A$ in a ring A, in particular, we also write a for $[a]_m$ in $\mathbb{Z}_m$.

simultaneous congruences

$$x \equiv a_1 \bmod m_1, \ldots, x \equiv a_r \bmod m_r$$

have a solution in $\mathbb{Z}$. *Moreover, a solution is uniquely determined modulo* m.

One can reformulate the Chinese Remainder Theorem in terms of ring homomorphisms as:

2.2 Corollary *Let* $m_1, \ldots, m_r$ *be pairwise relatively prime positive integers and let* $m := m_1 \cdots m_r$. *Then the canonical ring homomorphism*

$$\overline{\chi} : \mathbb{Z}_m \xrightarrow{\sim} \mathbb{Z}_{m_1} \times \cdots \times \mathbb{Z}_{m_r}, \quad [a]_m \mapsto ([a]_{m_1}, \ldots, [a]_{m_r}),$$

is an isomorphism.

In this formulation the Chinese Remainder Theorem is rather obvious: For arbitrary positive integers $m_1, \ldots, m_r$, the characteristic of the product ring $\mathbb{Z}_{m_1} \times \cdots \times \mathbb{Z}_{m_r}$ is the order

$$\operatorname{ord}([1]_{m_1}, \ldots, [1]_{m_r}) = \operatorname{LCM}(\operatorname{ord}[1]_{m_1}, \ldots, \operatorname{ord}[1]_{m_r}) = \operatorname{LCM}(m_1, \ldots, m_r).$$

In case the integers $m_1, \ldots, m_r$ are pairwise relatively prime, their LCM is the product m and, since also $\#(\mathbb{Z}_{m_1} \times \cdots \times \mathbb{Z}_{m_r}) = m$, the homomorphism $\overline{\chi}$ has to be bijective. For computational reasons it is important to describe the inverse of the isomorphism $\overline{\chi}$ of Corollary 2.2 conveniently.[9] This can be done by using the Euclidean algorithm in the following way: Let $n_\rho := m/m_\rho$, $\rho = 1, \ldots, r$. Then $\operatorname{GCD}(n_1, \ldots, n_r) = 1$ and there exists a representation of the unit $1 = b_1 n_1 + \cdots + b_r n_r$ with $b_1, \ldots, b_r \in \mathbb{Z}$ and the inverse of $\overline{\chi}$ is the map

$$([a_1]_{m_1}, \ldots, [a_r]_{m_r}) \mapsto [a_1 b_1 n_1 + \cdots + a_r b_r n_r]_m.$$

The Chinese Remainder Theorem can be formulated and proved without using any ring structure in the following way: *The product group* $G_1 \times \cdots \times G_r$ *of finite groups is cyclic if and only if every factor* $G_1, \ldots, G_r$ *is cyclic and the orders* $\#G_1, \ldots, \#G_r$ *are pairwise relatively prime.* This is (by induction on r) an immediate consequence of the following simple observation: If g, h are commuting elements of positive orders in a group G, then $\operatorname{ord} gh = \operatorname{ord} g \cdot \operatorname{ord} h$ if and only if $\operatorname{ord} g$ and $\operatorname{ord} h$ are relatively prime.

The isomorphism $\overline{\chi}$ of Corollary 2.2 induces an isomorphism of the groups of units:

2.3 Corollary *Let* $m_1, \ldots, m_r$ *be pairwise relatively prime positive integers and let* $m := m_1 \cdots m_r$. *Then the canonical group homomorphism*

$$\overline{\chi}^\times : \mathbb{Z}_m^\times \xrightarrow{\sim} \mathbb{Z}_{m_1}^\times \times \cdots \times \mathbb{Z}_{m_r}^\times$$

is an isomorphism. In particular, $\varphi(m) = \varphi(m_1) \cdots \varphi(m_r)$.

The finest decomposition of a positive integer m into pairwise relatively prime factors is the prime factorization $m := p_1^{\alpha_1} \cdots p_r^{\alpha_r}$ with pairwise distinct primes $p_1, \ldots, p_r$ and

[9] To give an example we mention the following method for the computation of the product of big integers b, c. If the absolute value of their product is $\leq n$, then one chooses (comparatively small) pairwise relatively prime positive integers (e. g. distinct prime numbers) $m_1, \ldots, m_r$ with $m := m_1 \cdots m_r \geq 2n+1$ and computes the product bc modulo each of the single numbers $m_1, \ldots, m_r$. Then by using Chinese Remainder Theorem we get the product bc modulo m, and hence bc itself, since $|bc| \leq n$.

positive exponents $\alpha_1, \ldots, \alpha_r$. From this it follows

$$\mathbb{Z}_m^\times \xrightarrow{\sim} \mathbb{Z}_{p_1^{\alpha_1}}^\times \times \cdots \times \mathbb{Z}_{p_r^{\alpha_r}}^\times$$

and

$$\varphi(m) = \varphi(p_1^{\alpha_1}) \cdots \varphi(p_r^{\alpha_r}) = p_1^{\alpha_1 - 1}(p_1 - 1) \cdots p_r^{\alpha_r - 1}(p_r - 1)\,.$$

Further, it follows that for a multiple n of m the canonical group homomorphism $\mathbb{Z}_n^\times \to \mathbb{Z}_m^\times$ induced by the canonical (surjective) ring homomorphism $\mathbb{Z}_n \to \mathbb{Z}_m$ is also surjective, since this is trivially true for prime powers $n = p^\beta$.

To determine the structure of the group of units $\mathbb{Z}_m^\times$ for $m \in \mathbb{N}^*$, it is sufficient to consider the case that $m = p^\alpha$ is a power of a prime number p. For $m = p$, the prime ring $\mathbb{Z}_p$ is a field and the group $\mathbb{Z}_p^\times$ is cyclic of order $\varphi(p) = p - 1$. This is a special case of the following general theorem:

2.4 Theorem *Any finite subgroup of the multiplicative group $K^\times$ of a field K is cyclic.*

To prove 2.4, let $G \subseteq K^\times$ be a finite subgroup of order $m = m_1 \cdots m_r$ with $m_\rho = p_\rho^{\alpha_\rho}$, $\alpha_\rho > 0$, $p_1, \ldots, p_r$ pairwise distinct prime numbers, and let $n_\rho := m/m_\rho$. Then $1 = b_1 n_1 + \cdots + b_r n_r$ for some integers $b_1, \ldots, b_r$. For every $x \in G$, one has $x = x^{b_1 n_1} \cdots x^{b_r n_r} = x_1 \cdots x_r$ with factors $x_\rho := x^{b_\rho n_\rho} \in G$ for which $x_\rho^{m_\rho} = x^{b_\rho m} = 1$, $\rho = 1, \ldots, r$. Since the polynomial $X^{m_\rho} - 1$ has at most m_ρ zeros in the field K, the group G contains at most and hence exactly m_ρ elements x with $x^{m_\rho} = 1$, otherwise G would contain less than $m_1 \cdots m_r = m$ elements. Moreover, since $X^{m_\rho/p_\rho} - 1$ has at most m_ρ/p_ρ solutions, there exists necessarily an element $y_\rho \in G$ with $y_\rho^{m_\rho} = 1$, but $y_\rho^{m_\rho/p_\rho} \neq 1$. Then $\operatorname{ord} y_\rho = m_\rho$ and $\operatorname{ord} y_1 \cdots y_r = m_1 \cdots m_r = m$, i. e. $y_1 \cdots y_r$ generates the group G. □

For a prime power p^α with $\alpha \geq 2$ and an odd prime p, we consider the exact sequence

$$1 \longrightarrow U \longrightarrow \mathbb{Z}_{p^\alpha}^\times \xrightarrow{\overline{\chi}^\times} \mathbb{Z}_p^\times \longrightarrow 1\,,$$

where U is the kernel of the surjective group homomorphism $\overline{\chi}^\times$ induced by the canonical ring homomorphism $\overline{\chi} : \mathbb{Z}_{p^\alpha} \to \mathbb{Z}_p$. The group U has order $p^{\alpha-1}$ and is cyclic with generator $1 + p$, since $(1+p)^{p^{\alpha-2}} \equiv 1 + p^{\alpha-1} \not\equiv 1 \bmod p^\alpha$, which follows directly by induction on α. Since $\overline{\chi}^\times$ is surjective and since $\mathbb{Z}_p^\times$ is cyclic of order $p - 1$, there is an element $z \in \mathbb{Z}_{p^\alpha}^\times$ of order $p - 1$, too. Then $(1+p)z$ has order $p^{\alpha-1}(p-1) = \#\mathbb{Z}_{p^\alpha}^\times$ and hence generates this group. Altogether:

2.5 Theorem *For an odd prime number p and any $\alpha \geq 1$, the group $\mathbb{Z}_{p^\alpha}^\times$ is cyclic of order $p^{\alpha-1}(p-1)$.*

The group $\mathbb{Z}_8^\times$ is isomorphic to $\mathbf{Z}_2 \times \mathbf{Z}_2$ and hence not cyclic. It follows that all the groups $\mathbb{Z}_{2^\alpha}^\times$ of order $2^{\alpha-1}$ are not cyclic for $\alpha \geq 3$. Indeed, as one can easily check by using $5^{2^{\alpha-3}} \equiv 1 + 2^{\alpha-1} \not\equiv \pm 1 \bmod 2^\alpha$ for $\alpha \geq 3$:

2.6 Theorem *For any $\alpha \geq 3$, the group $\mathbb{Z}_{2^\alpha}^\times$ is isomorphic to $\mathbf{Z}_{2^{\alpha-2}} \times \mathbf{Z}_2$. More precisely, the residue class $[5]$ is an element of order $2^{\alpha-2}$ in $\mathbb{Z}_{2^\alpha}^\times$ and $-[1]$ is an element of order 2 which does not belong to the subgroup generated by $[5]$.*

For an arbitrary $m = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$, $p_1 < \cdots < p_r$ primes, $\alpha_1, \ldots, \alpha_r > 0$, in the decomposition $\mathbb{Z}_m^\times \cong \mathbb{Z}_{p_1^{\alpha_1}}^\times \times \cdots \times \mathbb{Z}_{p_r^{\alpha_r}}^\times$ according to Corollary 2.3, all non-trivial factors are of even order, hence $\mathbb{Z}_m^\times$ is not cyclic if at least two such factors occur. It follows:

2.7 Theorem (Gauss) *Let $m \in \mathbb{N}^*$. Then the group $\mathbb{Z}_m^\times$ is cyclic if and only if m is of the form $m = 1,\ 2,\ 4,\ p^\alpha,\ 2p^\alpha$, where p is an arbitrary odd prime number and $\alpha \in \mathbb{N}^*$ is arbitrary.*

If $\mathbb{Z}_m^\times$ is cyclic, then any generator of the group $\mathbb{Z}_m^\times$ is called a primitive prime residue class modulo m.

2.8 Example For a given $m \in \mathbb{N}^*$ and a given $a \in \mathbb{Z}$, it is easy to decide whether $[a]_m$ belongs to the group $\mathbb{Z}_m^\times$ of units in $\mathbb{Z}_m$: one has to compute $\mathrm{GCD}(a,m)$. This is done by using Euclidean algorithm, which even yields a representation

$$d := \mathrm{GCD}(a,m) = ab + mr \quad \text{with} \quad b, r \in \mathbb{Z}.$$

Then $[a]_m$ is a unit if and only if $d = 1$ and, in this case, the equation $[1]_m = [a]_m[b]_m$ shows that $[b]_m = [a]_m^{-1}$ is the inverse of $[a]_m$ in $\mathbb{Z}_m^\times$. Now, *assume* $[a]_m \in \mathbb{Z}_m^\times$. To compute the order

$$\mathrm{ord}_m a := \mathrm{ord}\,[a]_m$$

of $[a]_m$ in the group $\mathbb{Z}_m^\times$, Euler's equation $[a]_m^{\varphi(m)} = [1]_m$ shows that $\mathrm{ord}_m a$ is a divisor of $\varphi(m)$ (and, by the way, $[a]_m^{-1} = [a]_m^{\varphi(m)-1}$). From the prime factorization $m = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$ of m one gets, by Corollary 2.3,

$$\varphi(m) = \varphi(p_1^{\alpha_1}) \cdots \varphi(p_r^{\alpha_r}) = p_1^{\alpha_1 - 1} \cdots p_r^{\alpha_r - 1}(p_1 - 1) \cdots (p_r - 1)$$

and

$$\mathrm{ord}_m a = \mathrm{LCM}(\mathrm{ord}_{p_1^{\alpha_1}} a, \ldots, \mathrm{ord}_{p_r^{\alpha_r}} a),$$

which reduces the problem to the case that $m = p^\alpha$ is a prime power > 1 and $\varphi(m) = p^{\alpha-1}(p-1)$. To assess the divisors of $\varphi(m)$ in this case, in addition the prime factorization of $p-1$ is required. If all these data are given, then we have an equation $[a]_m^n = 1$ for an $n \in \mathbb{N}^*$ with known prime factorization $n = q_1^{\beta_1} \cdots q_s^{\beta_s}$. In this situation the order of $[a]_m$ is rather easy to compute.

The following method can be used in any group G. The powers a^t, $t \in \mathbb{N}$, of an element $a \in G$, are quickly calculated by the method of rapid powers.[10] Now, assume that $a^n = e_G$ with n as above and let $n_\sigma := n/q_\sigma^{\beta_\sigma}$, $\sigma = 1, \ldots, s$. *If $\gamma_\sigma\ (\leq \beta_\sigma)$ is the smallest non-negative integer with $(a^{n_\sigma})^{q_\sigma^{\gamma_\sigma}} = e_G$, $\sigma = 1, \ldots, s$, then the order of a is*

$$\mathrm{ord}\, a = q_1^{\gamma_1} \cdots q_s^{\gamma_s}.$$

For the proof observe that, by construction, for all $\sigma = 1, \ldots, s$, $\mathrm{ord}\, a$ divides $n_\sigma q_\sigma^{\gamma_\sigma}$ and hence also divides $\mathrm{GCD}(n_1 q_1^{\gamma_1}, \ldots, n_s q_s^{\gamma_s}) = q_1^{\gamma_1} \cdots q_s^{\gamma_s}$. On the other hand, for every proper divisor t of $q_1^{\gamma_1} \cdots q_s^{\gamma_s}$ one has $a^t \neq e_G$ by the minimality of γ_σ, $\sigma = 1, \ldots, s$.

The inverse of the exponentiation map $a \mapsto a^t$ in the group G leads to discrete logarithms. To describe them, let y be a further element of G. The discrete logarithm problem (DLP) for the data $(G; a, y)$ asks for the existence of an exponent $x \in \mathbb{Z}$ with $y = a^x$. Moreover, if x exists

[10] Let $t = \sum_{i=0}^{\ell} t_i 2^i$ with $t_i \in \{0,1\}$ be the dual-expansion of t. Compute $a_i := a^{2^i}$ recursively by $a_0 = a$, $a_{i+1} = a_i^2$. Then a^t is the product of those a_i for which $t_i = 1$. – The power a^t is also the last element b_0 in the sequence $b_{\ell+1}, \ldots, b_0$ recursively constructed by $b_{\ell+1} = 1$, $b_i^2 = b_{i+1}^2 a^{t_i}$ for $i = \ell, \ldots, 0$.

one should compute such an exponent. If a solution x exists it is unique only modulo the order of a. In particular, if $n := \operatorname{ord} a$ is positive, then there is a smallest $x \in \mathbb{N}$ with $x < n$ and $a^x = y$ if the given DLP is solvable. This x is called *the* logarithm of y with respect to the base a and is denoted by

$$\log_a y .$$

In general a DLP is considered to be a difficult problem. But, using an idea of Pohlig and Hellmann, it is rather easy to solve if $n := \operatorname{ord} a > 0$ is known and if, moreover, a decomposition $n = n_1 \cdots n_r$ into "small" (not necessarily distinct) factors $n_1, \ldots, n_r \in \mathbb{N}^*$ is given, for instance, if the prime factors $q_1, \ldots, q_s$ from above are "small".[11] For the proof of this, one writes the potential solution

$$x = x_0 + x_1 n_1 + x_2 n_1 n_2 + \cdots + x_{r-1} n_1 \cdots n_{r-1}$$

with "digits" $x_0, \ldots, x_{r-1}$, $0 \le x_0 < n_1, \ldots, 0 \le x_{r-1} < n_r$. To compute x_0, one has to solve the equation

$$y = a^{x_0} a^{x_1 n_1 + \cdots + x_{r-1} n_1 \cdots n_{r-1}} \quad \text{or} \quad y^{n_2 \cdots n_r} = (a^{n_2 \cdots n_r})^{x_0} .$$

(Note that $\operatorname{ord} a = n_1 (n_2 \cdots n_r)$.) Since $\operatorname{ord} a^{n_2 \cdots n_r} = n_1$ and since n_1 is "small" by assumption this DLP can be solved by checking step by step computing successively the powers of $a^{n_2 \cdots n_r}$ (or by any other method). If there is no solution, then the original DLP has no solution. Otherwise, we have now to solve the equation

$$y a^{-x_0} = (a^{n_1})^{x_1 + x_2 n_2 + \cdots + x_{r-1} n_2 \cdots n_{r-1}} ,$$

which is the DLP for the data $(G; a^{n_1}, y a^{-x_0})$ with $\operatorname{ord} a^{n_1} = n_2 \cdots n_r$. Now, one proceeds for the computation of x_1 by solving a DLP with data $(G; a^{n_1 n_3 \cdots n_{r-1}}, (y a^{-x_0})^{n_3 \cdots n_{r-1}})$.

The discrete logarithm problems for the groups $\mathbb{Z}_m^\times$ are already difficult enough to use them in cryptography and security systems. For instance, one takes $m = p$ a prime number for which $p - 1$ has a "huge" prime factor q and for the base $a \in \mathbb{Z}_p^\times$ of the discrete logarithm an element of order q. – In this connection the Sophie Germain pairs $(q, p = 2q + 1)$ of primes q, p are of interest.[12]

The power maps for the groups $\mathbb{Z}_m^\times$ are also involved in the RSA-codes.[13] The background of these codes is the following simple observation: Let G be any finite group of order n. For any integer $r \in \mathbb{Z}$, the power map $G \to G$, $x \mapsto x^r$, is bijective if and only if $\operatorname{GCD}(r, n) = 1$.[14] In this case the inverse is also a power map. More precisely, the power map $x \mapsto x^s$ is inverse to $x \mapsto x^r$ if and only if $x^{rs} = x$ or, equivalently, $x^{rs-1} = e_G$ for every $x \in G$, i. e. $rs \equiv 1 \bmod \operatorname{Exp} G$.[15] Now, a (secret) message is interpreted as an element $x \in G$ and encoded as $y = x^r$ with $r \in \mathbb{N}^*$ relatively prime to $n := \# G$. To recover x from y one needs the inverse $[s]$ of $[r]$ in the group $\mathbb{Z}_{\operatorname{Exp} G}^\times$. Of course, it would be enough to find the inverse of $[r]$ in $\mathbb{Z}_n^\times$. This gives a cryptosystem if one can compute the powers in G easily without knowing $\# G$ or $\operatorname{Exp} G$. Then the enciphering

---

[11] As before one assumes that the multiplication in G and, in particular, exponentiation can be performed easily and, in addition, that the memory of the computer allows to solve DLP's $(G; b, z)$ for elements b of orders $n_1, \ldots, n_r$ in a reasonable time.

[12] For instance, $(q, 2q+1)$ with $q := 183\,027 \cdot 2^{265\,440} - 1$ is such a pair of primes found in March 2010.

[13] These codes were proposed by R. Rivest, A. Shamir and L. Adleman in 1977.

[14] This follows from the Theorem of Cauchy, see Example 1.5 (3). For an abelian group this is very simple to prove. As already remarked, the general case is also a consequence of the Theorem 1.6.1 (i) of Sylow.

[15] The exponent $\operatorname{Exp} G$ of a group G is, by definition, the unique non-negative generator of the subgroup $\{t \in \mathbb{Z} \mid x^t = e_G \text{ for all } x \in G\} \subseteq \mathbb{Z}$. If G is finite, then $\operatorname{Exp} G$ divides $\# G$ and, by the Theorem of Cauchy, $\operatorname{Exp} G$ and $\# G$ have the same prime divisors. If G is finite *and abelian*, then there exists an element $g \in G$ with $\operatorname{ord} g = \operatorname{Exp} G$.

exponent r (together with the group G) is the public key and the private deciphering exponent s is only known to those who know $\#G$ or $\operatorname{Exp} G$.

In the special case of RSA-codes the group G is $\mathbb{Z}_m^\times$, where $m = pq$ is a product of two large distinct primes p,q (of comparable order). The knowledge of the order $\#\mathbb{Z}_m^\times = (p-1)(q-1)$ is equivalent with the knowledge of the factors p,q of m.[16] For instance, one might pick distinct random primes p and q with about 200 digits each, so that m has roughly 400 digits. Its security depends on the assumption that in the current state of computer technology, the factorization of composite numbers with large prime factors is prohibitively time consuming. For computations in $\mathbb{Z}_m^\times$ only the knowledge of m is required, so the public key consists of a pair $(m = pq, r)$, where $r \in \mathbb{N}^*$ is relatively prime to $\varphi(m) = (p-1)(q-1)$ and the message to encode is given by an $x \in \mathbb{N}^*$ with $0 < x < m$ and $\operatorname{GCD}(x,m) = 1$.[17]

2.9 Example (Affine group of $\mathbb{Z}_m$) Let $m \in \mathbb{N}$. The group $\mathbb{Z}_m^\times$ of units operates canonically by group automorphisms on the additive group $\mathbb{Z}_m$ of the ring $\mathbb{Z}_m$ by multiplication $\mathbb{Z}_m^\times \times \mathbb{Z}_m \to \mathbb{Z}_m$, $([b],[x]) \mapsto [b][x] = [bx]$. The action homomorphism $\vartheta : \mathbb{Z}_m^\times \to \operatorname{Aut} \mathbb{Z}_m$ is obviously bijective. Hence, *the automorphism group of a cyclic group $\mathbf{Z}_m$ can be identified canonically with the group $\mathbb{Z}_m^\times$ and the semidirect product $\mathbf{Z}_m \rtimes \mathbb{Z}_m^\times$ is the full holomorph* $\operatorname{Hol}(\mathbf{Z}_m)$ *of* $\mathbf{Z}_m$. The group $\mathbb{Z}_m \rtimes \mathbb{Z}_m^\times (= \operatorname{Hol}(\mathbb{Z}_m))$ is also called the affine group of the ring $\mathbb{Z}_m$.[18] An element $([a],[b]) \in \mathbb{Z}_m \rtimes \mathbb{Z}_m^\times$, $a,b \in \mathbb{Z}$, $\operatorname{GCD}(b,m) = 1$, is identified with the affine transformation

$$([a],[b]) = (a,b) = (a,b)_m : \mathbb{Z}_m \to \mathbb{Z}_m, \quad x \mapsto a + bx.$$

(Recall the convention of Footnote 8.) For $b \in \mathbb{Z}_m^\times$, $(0,b)_m = \vartheta_b$ is the homothecy with b, i. e. the multiplication by b in $\mathbb{Z}_m$. The homothecies are exactly those affine transformations which have $0 \in \mathbb{Z}_m$ as a fixed point. 0 is the only fixed point of ϑ_b if and only if besides b also $1-b$ is a unit in $\mathbb{Z}_m$. Moreover, in this case every transformation $(a,b)_m$ has exactly one fixed point, namely $a/(1-b)$. If x_0 is any fixed point of $(a,b)_m$ (not necessarily the only one), then $(a,b)_m = (x_0,1)_m (0,b)_m (-x_0,1)_m$ is a conjugate of the homothecy $\vartheta_b = (0,b)_m$ and, in particular, $\operatorname{ord}(a,b)_m = \operatorname{ord}(0,b)_m = \operatorname{ord}_m b$.

For $m = 0$, the group $\mathbb{Z} \rtimes \mathbb{Z}^\times$ is isomorphic to the infinite dihedral group $\mathbf{D}_0 = \operatorname{Hol}(\mathbf{Z}_0)$ and is generated, for example, by the two reflections $(1,-1)$ and $(0,-1)$, the product of which is the translation $(1,1) = (x \mapsto 1+x)$ of order 0.

Now, let $m \in \mathbb{N}^*$. Then the orbits of the canonical action $\mathbb{Z}_m^\times \times \mathbb{Z}_m \to \mathbb{Z}_m$ are precisely the subsets $X_d := \{x \in \mathbb{Z}_m \mid \operatorname{ord}_{\mathbb{Z}_m} x = d\}$ of cardinality $\varphi(d)$, where d is a divisor of m.[19] This follows directly from the surjectivity of the canonical group homomorphism $\mathbb{Z}_m^\times \to \mathbb{Z}_d^\times$ induced by the ring homomorphism $\mathbb{Z}_m \to \mathbb{Z}_d$, $d|m$, see the remark after Corollary 2.3.

It follows from the Chinese Remainder Theorem 2.1 that for $m = m_1 \cdots m_r$ with pairwise rela-

[16] Note that $\operatorname{Exp} \mathbb{Z}_m^\times = \operatorname{LCM}(p-1, q-1)$. Thus, in choosing p,q one should also think of $\operatorname{GCD}(p-1,q-1)$.

[17] If one chooses incidentally a message x with $0 < x < m$ and $\operatorname{GCD}(x,m) \neq 1$, then $\operatorname{GCD}(x,m) \in \{p,q\}$ and the factorization of m would be known which makes the code useless. But, this happens very rarely with probability $(p+q-2)/(m-1)$ and is, moreover, irrelevant for the algorithm itself, because the power map $x \mapsto x^r$ is also a permutation of $\mathbb{Z}_m$ (and not only of $\mathbb{Z}_m^\times$) with inverse $x \mapsto x^s$. – If one knows the order $t := \operatorname{ord}_m x^r = \operatorname{ord}_m x$ of x^r or x in $\mathbb{Z}_m^\times$ then one also knows the message, namely $x = (x^r)^u$, where $ru \equiv 1 \bmod t$. If, moreover, t is even and $x^{t/2} \not\equiv -1 \bmod m$, then $0 = x^t - 1 = (x^{t/2}-1)(x^{t/2}+1)$ in $\mathbb{Z}_m$ and hence $\operatorname{GCD}(x^{t/2}-1, m) \in \{p,q\}$ provides now the factorization of m. Quantum computers will allow to compute t in a reasonable time (in the future).

[18] More generally, for an arbitrary ring A, the affine group of A is the group $A \rtimes A^\times = \operatorname{Hol}_{A^\times}(A)$, where $A^\times$ operates canonically on the additive group of A by left multiplication.

[19] In this case the class equation yields the well known formula $m = \#\mathbb{Z}_m = \sum_{d|m} \# X_d = \sum_{d|m} \varphi(d)$.

tively prime positive integers $m_1,\dots,m_r$, the canonical group homomorphism

$$\mathbb{Z}_m \rtimes \mathbb{Z}_m^\times \longrightarrow (\mathbb{Z}_{m_1} \rtimes \mathbb{Z}_{m_1}^\times) \times \cdots \times (\mathbb{Z}_{m_r} \rtimes \mathbb{Z}_{m_r}^\times), \quad (a,b)_m \mapsto ((a,b)_{m_1},\dots,(a,b)_{m_r})$$

is an isomorphism

The Linear Congruence Method for generating "random numbers" uses the affine transformations $(a,b)_m$: Starting with an arbitrary $x_0 \in \mathbb{Z}_m$, one forms the sequence (x_i) in $\mathbb{Z}_m$ by the recursion

$$x_{i+1} = (a,b)_m(x_i) = a + bx_i, \quad i \in \mathbb{N}.$$

This sequence is purely periodic of a length t which divides $\operatorname{ord}(a,b)_m$. The cycle $\langle x_0, x_1, \dots, x_{t-1}\rangle$ belongs to the cycle decomposition of $(a,b)_m \in \mathfrak{S}(\mathbb{Z}_m)$, cf. Example 1.7. For $x_0 = 0$, one obtains the sequence

$$x_i = a(1 + b + \cdots + b^{i-1}), \quad i \in \mathbb{N}.$$

This is

$$x_i = a \cdot \frac{1-b^i}{1-b}, \quad i \in \mathbb{N},$$

if, in addition, $1 - b \in \mathbb{Z}_m^\times$ (which, by the way, is possible only for odd m). If, moreover, also $a = x_1 \in \mathbb{Z}_m^\times$, then $x_i = 0$ for exactly $i \in \mathbb{N} \cdot \operatorname{ord}_m b$, i. e. the sequence (x_i) is purely periodic of period length $\operatorname{ord}_m b = \operatorname{ord}(a,b)_m$. For example, if $m = p$ is an odd prime and if $0 < a, b < p$, where b is a primitive prime residue class modulo p (i. e. $\operatorname{ord}_p b = p-1$), then, starting with an arbitrary x_0, $0 \le x_0 < p$, which is not the fixed point $a/(1-b)\,(\neq 0)$ of $(a,b)_p$, the terms of the sequence (x_i) run periodically through all the elements of $\mathbb{Z}_p$ with the exception of the fixed point $a/(1-b)$.

The linear congruence method for generating random numbers is discussed more thoroughly in [4, 3.2.1].

2.10 Generalized Chinese Remainder Theorem Let $m_1,\dots,m_r$ be positive integers and let $m := \operatorname{LCM}(m_1,\dots,m_r)$. The Generalized Chinese Remainder Theorem characterizes the elements of the image of the canonical homomorphism $\chi : \mathbb{Z} \to \mathbb{Z}_{m_1} \times \cdots \times \mathbb{Z}_{m_r}$ also in the case that the integers $m_1,\dots,m_r$ are not pairwise relatively prime. Further, note that $([a]_{m_1},\dots,[a]_{m_1}) \in \operatorname{im}\chi$ if and only if the system $x \equiv a_1 \bmod m_1, \dots, x \equiv a_r \bmod m_r$ of simultaneous congruences has a solution in $\mathbb{Z}$. More precisely, we have the following result:

2.10.1 Generalized Chinese Remainder Theorem *In the situation as above, for a given r-tuple of integers $(a_1,\dots,a_r) \in \mathbb{Z}^r$ the simultaneous congruences*

$$x \equiv a_1 \bmod m_1, \dots, x \equiv a_r \bmod m_r$$

have a solution in $\mathbb{Z}$ if and only if the congruences $a_i \equiv a_j \bmod \operatorname{GCD}(m_i,m_j)$ hold for all pairs (i,j), $1 \le i,j \le r$. Moreover, a solution is uniquely determined modulo $m = \operatorname{LCM}(m_1,\dots,m_r)$.

To prove 2.10.1, let x be a solution. Then $x - a_i \in \mathbb{Z}m_i$, $x - a_j \in \mathbb{Z}m_j$ and $a_i - a_j = -(x-a_i) + (x-a_j) \in \mathbb{Z}m_i + \mathbb{Z}m_j = \mathbb{Z}\operatorname{GCD}(m_i,m_j)$. For the converse we proceed by induction on r. The cases $r = 1$ and $r = 2$ are trivial: if $a_1 \equiv a_2 \bmod d := \operatorname{GCD}(m_1,m_2)$, i. e. $a_1 - a_2 \in \mathbb{Z}m_1 + \mathbb{Z}m_2$, $a_1 = a_2 + b_1m_1 + b_2m_2$, then $a := a_1 - b_1m_1 = a_2 + b_2m_2$ is a solution of the congruences $x \equiv a_i \bmod m_i$, $i = 1,2$. Now, let $r > 2$. Then let $a' \in \mathbb{Z}$ be

a solution of the $r-1$ congruences $a' \equiv a_i \bmod m_i$, $i = 1,\dots,r-1$. Any solution of the two congruences $x \equiv a' \bmod \mathrm{LCM}(m_1,\dots,m_{r-1}) =: m'$, $x \equiv a_r \bmod m_r$ will be a solution of the given system of r congruences. That the last two congruences have a common solution follows from $a'-a_r = (a'-a_i)+(a_i-a_r) \in \mathbb{Z}m_i + \mathbb{Z}m_r$ for $i = 1,\dots,r-1$, that is, $a'-a_r \in \bigcap_{i=1}^{r-1}(\mathbb{Z}m_i+\mathbb{Z}m_r) = \left(\bigcap_{i=1}^{r-1}\mathbb{Z}m_i\right)+\mathbb{Z}m_r = \mathbb{Z}m'+\mathbb{Z}m_r$. Here we have used the distributive law $(\mathfrak{a}+\mathfrak{c})\cap(\mathfrak{b}+\mathfrak{c}) = \mathfrak{a}\cap\mathfrak{b}+\mathfrak{c}$ for ideals $\mathfrak{a} = \mathbb{Z}a, \mathfrak{b} = \mathbb{Z}b, \mathfrak{c} = \mathbb{Z}c$ in the ring $\mathbb{Z}$. (By the way, also the distributive law $\mathfrak{a}\cap\mathfrak{c}+\mathfrak{b}\cap\mathfrak{c} = (\mathfrak{a}+\mathfrak{b})\cap\mathfrak{c}$ holds for arbitrary ideals in $\mathbb{Z}$. For the generators $a,b,c \in \mathbb{Z}$ of these ideals the distributive laws are the well-known (and simple) formulae: $\mathrm{LCM}(\mathrm{GCD}(a,c),\mathrm{GCD}(b,c)) = \mathrm{GCD}(\mathrm{LCM}(a,b),c)$ and $\mathrm{GCD}(\mathrm{LCM}(a,c),\mathrm{LCM}(b,c)) = \mathrm{LCM}(\mathrm{GCD}(a,b),c)$.)

To compute a solution in the case of Theorem 2.10.1 one may proceed in a similar manner as in the special case of Theorem 2.1. Now, one can construct a projection homomorphism (of the additive groups)

$$\varphi : \mathbb{Z}_{m_1} \times \cdots \times \mathbb{Z}_{m_r} \to \mathbb{Z}_m$$

with $\varphi\overline{\chi} = \mathrm{id}_{\mathbb{Z}_m}$, where $\overline{\chi} : \mathbb{Z}_m \to \mathbb{Z}_{m_1} \times \cdots \times \mathbb{Z}_{m_r}$ is the *injective* homomorphism induced by χ and $m = \mathrm{LCM}(m_1,\dots,m_r)$. Because of $\mathrm{GCD}(n_1,\dots,n_r) = 1$ for $n_\rho := m/m_\rho$, $\rho = 1,\dots,r$, there exist $b_1,\dots,b_r \in \mathbb{Z}$ with $1 = b_1n_1 + \cdots + b_rn_r$ and $\varphi([a_1]_{m_1},\dots,[a_r]_{m_r}) = [a_1b_1n_1 + \cdots + a_rb_rn_r]_m$ is well-defined and does the job. We get:

2.10.2 Proposition *Let $m_1,\dots,m_r \in \mathbb{N}^*$ be positive integers, $m := \mathrm{LCM}(m_1,\dots,m_r)$, $n_1 := m/m_1,\dots,n_r := m/m_r$ and $1 = b_1n_1 + \cdots + b_rn_r$ with $b_1,\dots,b_r \in \mathbb{Z}$. Then the simultaneous congruences*

$$x \equiv a_1 \bmod m_1,\dots,x \equiv a_r \bmod m_r$$

have a solution if and only if

$$a \equiv a_1 \bmod m_1,\dots,a \equiv a_r \bmod m_r \quad \text{for } a := a_1b_1n_1 + \cdots + a_rb_rn_r.$$

In this case a is the unique solution modulo m.

It is very easy to show that in 2.10.2 the congruences $a_i \equiv a_j \bmod \mathrm{GCD}(m_i,m_j)$ for $1 \le i < j \le r$ imply the congruences $a \equiv a_i \bmod m_i$ for $1 \le i \le r$. Therefore, 2.10.2 provides an independent (and constructive) proof of 2.10.1.

2.11 Quadratic Residues In the next section very often we have to compute the order $\mathrm{ord}_m 2$ of $[2]_m \in \mathbb{Z}_m^\times$ for an odd integer $m \in \mathbb{N}^*$. This is always a divisor of $\varphi(m) = \#\mathbb{Z}_m^\times$. But, if $[2]_m = [a]_m^2$ is a square in $\mathbb{Z}_m^\times$ and if $m > 1$, i. e. $2 \mid \varphi(m)$, then $\mathrm{ord}_m 2$ divides even $\varphi(m)/2$ because of $[2]_m^{\varphi(m)/2} = [a]_m^{\varphi(m)} = [1]_m$. If $\mathbb{Z}_m^\times$ is cyclic, i. e. if $m = p^\alpha$ is an odd prime power > 1 (cf. Theorem 2.7), then the converse is true: If $2^{\varphi(m)/2} \equiv 1 \bmod m$ then 2 is a square in $\mathbb{Z}_m^\times$. More generally, the following simple lemma holds:

2.11.1 Lemma *Let G be a finite cyclic group of order $n \in \mathbb{N}^*$ and let $r \in \mathbb{Z}$. An element $g \in G$ is an r-th power in G if and only if $g^{n/\mathrm{GCD}(n,r)} = e_G$,* i. e. *if and only if* ord g *divides $n/\mathrm{GCD}(n,r)$.*

For the proof, let $a \in G$ be a generator of G. Then the subgroup of the r-th powers in G is generated by a^r and hence of order $\operatorname{ord} a^r = s := n/\mathrm{GCD}(n,r)$. But this subgroup is the kernel of the power endomorphism $x \mapsto x^s$ of G.

The theory of squares in $\mathbb{Z}_m^\times$ is the theory of quadratic residues:

2.11.2 Definition Let $m \in \mathbb{N}^*$ and $a \in \mathbb{Z}$ with $\mathrm{GCD}(m,a) = 1$. Then a is called a quadratic residue modulo m if there exists a $b \in \mathbb{Z}$ with $a \equiv b^2 \bmod m$, i. e. $[a]_m$ is a square in the group $\mathbb{Z}_m^\times$.

Sometimes the condition $\mathrm{GCD}(m,a) = 1$ in Definition 2.11.2 is omitted. But, here we keep it as we want to focus on squares in $\mathbb{Z}_m^\times$. By the Chinese Remainder Theorem (cf. Corollary 2.3) the integer a is a quadratic residue modulo m if and only if a is a quadratic residue modulo all (maximal) prime powers p^α dividing m. Furthermore, if p is an odd prime number then a is a quadratic residue modulo p^α, $\alpha \geq 1$, if and only if a is a quadratic residue modulo p. This follows from the exact sequence after the proof of Theorem 2.4, which was used for the proof of Theorem 2.5. For $\alpha \geq 3$, an integer a is a quadratic residue modulo 2^α if and only if it is a quadratic residue modulo $8 = 2^3$, i. e. if and only if $a \equiv 1 \bmod 8$ (since the group of squares of $\mathbb{Z}_{2^\alpha}^\times$ belongs to the kernel of $\mathbb{Z}_{2^\alpha}^\times \to \mathbb{Z}_8^\times$ and has the same index 4 as this kernel, by Theorem 2.6, for example).

From now on we consider only quadratic residues modulo p where p is an odd prime.

For describing the quadratic residues modulo p it is convenient to introduce the Legendre symbol

$$\left(\frac{a}{p}\right) = (a/p) := \begin{cases} 1, & \text{if } [a]_p \text{ is a square in } \mathbb{Z}_p^\times, \\ -1, & \text{if } [a]_p \text{ is not a square in } \mathbb{Z}_p^\times, \end{cases}$$

for an integer a not divisible by p. Since the image of the group homomorphism $[a]_p \mapsto [a]_p^{(p-1)/2}$ in $\mathbb{Z}_p^\times$ is the subgroup $\{\pm 1\} \subseteq \mathbb{Z}_p^\times$ of order 2, by Lemma 2.11.1, this homomorphism coincides with the Legendre symbol $[a]_p \mapsto (a/p)$, in particular $(ab/p) = (a/p)(b/p)$ for all integers a, b which are not divisible by p. For later use we note explicitly:

2.11.3 Euler's Criterion for Quadratic Residues *For an odd prime p and an integer a not divisible by p, we have*

$$\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} \bmod p.$$

Note that $[a]_p \mapsto (a/p)$ is the only non-trivial group homomorphism $\mathbb{Z}_p^\times \to \{\pm 1\}$. Since $[a]_p \mapsto \operatorname{sign} \mathrm{L}_a$ is also such a homomorphism (where L_a denotes multiplication with a in $\mathbb{Z}_p$ or $\mathbb{Z}_p^\times$) we have

$$(a/p) = \operatorname{sign} \mathrm{L}_a$$

for an integer a not divisible by p.[20]

[20] In this way the Legendre symbol may be generalized to arbitrary integers a, m with m odd and $\mathrm{GCD}(a,m) = 1$ by setting

As a special case of 2.11.3 we get $(-1/p) = (-1)^{(p-1)/2}$ and hence the following:

2.11.4 Lemma *The integer -1 is a quadratic residue modulo an odd prime p if and only if $p \equiv 1 \bmod 4$,* i. e. $(-1/p) \equiv p \bmod 4$.

Explicitly: If $p \equiv 1 \bmod 4$, then $\pm((p-1)/2)!$ are the solutions of the equation $x^2 = -1$ in $\mathbb{Z}_p$. (This follows from Wilson's Theorem: $(p-1)! \equiv -1 \bmod p$.)

The assertion 2.11.4 is the first supplement of the famous Quadratic Reciprocity Law: *For odd primes p, q, $p \neq q$,*

$$\left(\frac{q}{p}\right)\left(\frac{p}{q}\right) = (-1)^{\frac{(p-1)(q-1)}{4}}.$$

This theorem was already used by L. Euler (1707-1783). A first complete proof was given by C. F. Gauss (1777-1855) in his Disquisitiones arithmeticae (cf. [3]). It allows to characterize, for a given positive odd integer $a = q_1 \cdots q_s \in \mathbb{N}^*$ (which may be assumed to be square free) with pairwise distinct odd prime factors $q_1, \ldots, q_s$, those primes $p \notin \{q_1, \ldots, q_s\}$ for which a is a quadratic residue modulo p, namely

$$\left(\frac{a}{p}\right) = (-1)^{\frac{p-1}{2}\sum_{\sigma=1}^{s}\frac{q_\sigma-1}{2}} \cdot \prod_{\sigma=1}^{s}\left(\frac{p}{q_\sigma}\right).$$

The second supplement of the Quadratic Reciprocity Law concerns the Legendre symbol $(2/p)$:

2.11.5 Proposition *For an odd prime number p,*

$$\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}} = \begin{cases} 1, & \text{if } p \equiv \pm 1 \bmod 8, \\ -1, & \text{if } p \equiv \pm 3 \bmod 8. \end{cases}$$

There is a simple proof of 2.11.5 proposed by V. A. Lebesgue (1791-1875). It counts the points of the unit circle $S^1 = \mathrm{S}^1(\mathbb{Z}_p) = \{(x,y) \in \mathbb{Z}_p^2 \mid x^2 + y^2 = 1\}$ over $\mathbb{Z}_p$. The stereographic projection maps $\mathrm{S}^1(\mathbb{Z}_p) \setminus \{(0,1)\}$ bijectively onto $\mathbb{Z}_p \setminus \{t \in \mathbb{Z}_p \mid t^2 = -1\}$ via $(x,y) \mapsto x/(1-y)$, and its inverse is $t \mapsto (2t/(t^2+1), (t^2-1)/(t^2+1))$. It follows that

$$\#S^1 = 1 + (p - (1 + (-1/p))) = p - (-1/p).$$

$(a/m) := \operatorname{sign} \mathrm{L}_a$, where L_a denotes multiplication with a in $\mathbb{Z}_m$ or in $\mathbb{Z}_m^\times$. But here these signs may differ. (For example, consider $m = p^2$ where p is an odd prime.) If one chooses *multiplication in* $\mathbb{Z}_m$, one gets the so-called Jacobi symbol. If $[a/m]$ denotes the alternative symbol defined with the multiplication in $\mathbb{Z}_m^\times$, then, for all odd $m \in \mathbb{N}^*$ and for all $a \in \mathbb{Z}$ with $\mathrm{GCD}(a,m) = 1$, it follows from the description of the orbits of the canonical action $\mathbb{Z}_m^\times \times \mathbb{Z}_m \to \mathbb{Z}_m$ in Example 2.9 (cf. also the previous Footnote 19) that $(a/m) = \prod_{d|m}[a/d]$. Therefore

$$\left[\frac{a}{m}\right] = \prod_{d|m}\left(\frac{a}{d}\right)^{\mu(m/d)} = \prod_{\substack{d|m, \\ m/d \text{ square-free}}}\left(\frac{a}{d}\right)$$

by the (multiplicative) Möbius Inversion Formula: If $f, g : \mathbb{N}^* \to G$ are two maps into any (multiplicatively written) abelian group G with $f(m) = \prod_{d|m} g(d)$, then $g(m) = \prod_{d|m} f(d)^{\mu(m/d)}$, where μ denotes the classical Möbius function defined by $\mu(n) := (-1)^r$ if $n = p_1 \cdots p_r$ is the product of distinct prime numbers $p_1, \ldots, p_r$ and $\mu(n) := 0$ otherwise. This function was introduced by A. F. Möbius (1790-1868) in 1832 and is important in number theory and combinatorics (where it is generalized extensively). For further properties of the Jacobi symbol see the end of this subsection.

On the other hand, on S^1 operates the square group $\mathbf{D}_4$ of order 8 generated by the two reflections $(x,y)\mapsto(x,-y)$ and $(x,y)\mapsto(y,x)$ at the x-axis $\{y=0\}$ and at the diagonal $\{x=y\}$ respectively. All orbits have cardinality 8 with the following exceptions: In any case the points on the coordinates axes form an orbit $\{(\pm1,0),(0,\pm1)\}$ of cardinality 4, and, *if 2 is a square in* $\mathbb{Z}_p$, i. e. *if* $(2/p)=1$, the points on the two diagonals $x=\pm y$ also form an orbit $\{(\pm1/\sqrt{2},\pm1/\sqrt{2})\}$ of cardinality 4. Altogether $\#S^1 = 8k+4+2(1+(2/p))$, where k is the number of orbits of cardinality 8. From the equality $p-(-1/p)=8k+4+2(1+(2/p))$ one gets (besides $p\equiv(-1/p) \bmod 4$ which is 2.11.4) the two congruences $p\equiv(-1/p) \bmod 8$ if $(2/p)=1$ and $p\equiv(-1/p)+4 \bmod 8$ if $(2/p)=-1$, and this is 2.11.5.

The method of this proof can also be used to prove the main part of the Quadratic Reciprocity Law. For this one counts the points of the unit sphere

$$S^{q-1}=\mathrm{S}^{q-1}(\mathbb{Z}_p):=\{(x_1,\ldots,x_q)\in\mathbb{Z}_p^q | x_1^2+\cdots+x_q^2=1\}$$

over $\mathbb{Z}_p$. Quite generally, one has the cardinality formula $\#S^{2n}=p^n(p^n+(-1/p)^n)$ for $n\in\mathbb{N}$, which one proves (perhaps) most easily by induction on n using the recursion

$$\#S^0=2\,,\quad \#S^{m+2}=p^{m+1}\,(p-(-1/p))+\#S^m\cdot(-1/p)\,p\,.$$

(To prove this formula, one counts the elements in the fibres of the projection

$$S^{m+2}\to\mathbb{Z}_p^{m+1},\ (x_1,\ldots,x_{m+1},x_{m+2},x_{m+3})\mapsto(x_1,\ldots,x_{m+1})\,.$$

For any point in $\mathbb{Z}_p^{m+1}$ which does not belong to S^m, the fibre contains $p-(-1/p)\,(=\#S^1)$ points, and the fibre over any point in S^m contains $(1+(-1/p))p-(-1/p)$ points.) In particular, $\#S^{q-1}=p^{(q-1)/2}\left(p^{(q-1)/2}+(-1)^{(p-1)(q-1)/4}\right)$. On the other hand, the group $\mathbf{Z}_q$ operates on S^{q-1} canonically by the reiterated cyclic permutation $\langle1,2,\ldots,q\rangle$. The fixed points are the constant tuples $(x,\ldots,x)$ with $qx^2=1$ or $q=(x^{-1})^2$ in $\mathbb{Z}_p$. Hence, the number of fixed points is $1+(q/p)$. All the non-constant orbits have q points each. Altogether, the class equation for this operation of $\mathbf{Z}_q$ on S^{q-1} is $\#S^{q-1}=1+(q/p)+kq$ with some $k\in\mathbb{N}$. Comparing both expressions for $\#S^{q-1}$ and using again Euler's criterion $p^{(q-1)/2}\equiv(p/q) \bmod q$ from 2.11.3 yields $(p/q)\left((p/q)+(-1)^{(p-1)(q-1)/4}\right)\equiv1+(q/p) \bmod q$, which is the Quadratic Reciprocity Law, see also [6].

It is possible to look at Lebesgue's proof of 2.11.5 in another way. The essential point for the proof is to look at the points of the unit circle which belong to the diagonals $x=\pm y$. In the complex plane these are the primitive 8-th roots of unity $(\pm1\pm i)/\sqrt{2}$. If we choose one of them as ζ_8, then $\zeta_8^4=-1$, ζ_8^2 is a primitive 4-th root of unity and $\zeta_8+\zeta_8^{-1}$ is a square root of 2, since $(\zeta_8+\zeta_8^{-1})^2=\zeta_8^2+2+\zeta_8^{-2}=2$ (which expresses the fact that $\sqrt{2}$ is the length of the diagonal in a square of side length 1). The same equations hold in an extension field K of $\mathbb{Z}_p$ which contains a primitive 8-th root of unity again denoted by ζ_8.[21] It follows: $(2/p)=1$ if and only if $\sqrt{2}=\zeta_8+\zeta_8^{-1}\in\mathbb{Z}_p$, i. e. if and only if

[21] Let $n\in\mathbb{N}^*$ be a positive integer relatively prime to p. A finite field $\mathbb{F}_{p^r}$ of cardinality p^r, $r\in\mathbb{N}^*$, contains a primitive n-th roots of unity ζ_n if and only if $n\,|\,\#\mathbb{F}_{p^r}^\times$, i. e. if and only if $p^r\equiv1 \bmod n$, i. e. if and only if r is a multiple of $\mathrm{ord}_n p$. It follows $[\mathbb{Z}_p[\zeta_n]:\mathbb{Z}_p]=\mathrm{ord}_n p$. For $n=8$ this order is 1 if $p\equiv1 \bmod 8$ and 2 if $p\not\equiv1 \bmod 8$, i. e. $\mathbb{F}_{p^2}$ contains always an 8-th primitive

$(\zeta_8+\zeta_8^{-1})^p=\zeta_8^p+\zeta_8^{-p}=\zeta_8+\zeta_8^{-1}$. For $p\equiv\pm1 \bmod 8$ one has $\zeta_8^p+\zeta_8^{-p}=\zeta_8+\zeta_8^{-1}$ and for $p\equiv\pm3 \bmod 8$ one has $\zeta_8^p+\zeta_8^{-p}=\zeta_8^3+\zeta_8^5=\zeta_8^4(\zeta_8^{-1}+\zeta_8)=-(\zeta_8+\zeta_8^{-1})$. This proves 2.11.5. Also this proof can be generalized to prove the main part of the Quadratic Reciprocity Law, now using an extension field K of $\mathbb{Z}_p$ containing a primitive q-th root of unity ζ_q. Already Gauss observed (see [3]) that K contains a square root of $(-1)^{(q-1)/2}q$. He gave even an explicit expression for such a root, namely

$$z:=\sum_{a=1}^{q-1}\left(\frac{a}{q}\right)\zeta_q^a\,.$$

The reader may check this as well as the equality $z^p=(p/q)z$ which proves the Quadratic Reciprocity Law.[22]

For example, if $q=3$, then $z=\zeta_3-\zeta_3^{-1}$, $z^2=\zeta_3^2+\zeta_3-2=-1-2=-3$ (i.e. $\zeta_3=(-1\pm\sqrt{-3})/2$) and

$$z^p=\zeta_3^p-\zeta_3^{-p}=\begin{cases}\zeta_3-\zeta_3^{-1}=z, & \text{if } p\equiv\ \ 1 \bmod 3\,,\\ \zeta_3^{-1}-\zeta_3=-z, & \text{if } p\equiv -1 \bmod 3\,,\end{cases}$$

i. e. $(-3/p)=(-1)^{(p-1)/2}(3/p)=(p/3)$.[23]

We recommend to treat in a similar way the case $q=5$. But, *to prove the general case* we proceed here a little bit differently (without using the element z from above): The discriminant of the polynomial $f:=X^q-1=\prod_{i=0}^{q-1}(X-\zeta_q^i)\in\mathbb{Z}_p[X]\subseteq K[X]$ is (by definition)

$$(-1)^{q(q-1)/2}\prod_{0\le i<j<q}(\zeta_q^j-\zeta_q^i)^2=\prod_{i\ne j}(\zeta_q^j-\zeta_q^i)=\prod_{j=0}^{q-1}f'(\zeta_q^j)=q^q\prod_{j=0}^{q-1}\zeta_q^{-j}=q^q$$

and yields the equation $(-1)^{(q-1)/2}q^q=V^2$ with the Vandermonde determinant

$$V:=\mathrm{V}(1,\zeta_q,\dots,\zeta_q^{q-1})=\mathrm{Det}(\zeta_q^{ij})_{0\le i,j<q}\,.$$[24]

We get $((-1)^{(q-1)/2}q/p)=(-1)^{(p-1)(q-1)/4}(q/p)=1$ if and only if $V\in\mathbb{Z}_p$ or if

$$V=V^p=\mathrm{V}(1,\zeta_q^p,\dots,\zeta_q^{p(q-1)})=\mathrm{sign}(\mathrm{L}_p)\cdot V=(p/q)V\,,$$

i.e. $(p/q)=1$. (L_p denotes multiplication with p in $\mathbb{Z}_q$.) In this proof we may replace q by an arbitrary positive odd integer b not divisible by p and get $(-1)^{(p-1)(b-1)/4}(b/p)=$

root of unity. Or: The degree of an algebraic element x in an extension field K of $\mathbb{Z}_p$ is the smallest $r>0$ with $x^{p^r}=x$. For $x=\zeta_n$ this gives the same characterization as above.

[22]To compute $z^2=\sum_{a,b}(ab/q)\zeta_q^{a+b}$ we collect the summands with a fixed exponent $a+b\equiv c \bmod q$, $c=0,\dots,q-1$. For $c\equiv0 \bmod q$, this gives $\sum_{a=1}^{q-1}(a(q-a)/q)=\sum_{a=1}^{q-1}(-1/q)=(-1)^{(q-1)/2}(q-1)$. For $c\not\equiv0 \bmod q$, we get $(-1)^{(q-1)/2}\,\zeta_q^c\sum_{1\le a<q,a\ne c}(a(a-c)/q)$. In $\mathbb{Z}_q$ we have $a(a-c)=a^2(1-c/a)$, and $1-c/a$ runs through all elements $\ne1$ in $\mathbb{Z}_q^\times$, hence $\sum_{1\le a<q,a\ne c}(a(a-c)/q)=-1$. Altogether, as asserted, $z^2=(-1)^{(q-1)/2}\left(q-1-\sum_{c=1}^{q-1}\zeta_q^c\right)=(-1)^{(q-1)/2}q$, since $\sum_{c=0}^{q-1}\zeta_q^c=(\zeta_q^q-1)/(\zeta_q-1)=0$. It follows $\left((-1)^{(q-1)/2}q/p\right)=(-1)^{(p-1)(q-1)/4}(q/p)=1$ if and only if $z=z^p=\sum_{a=1}^{q-1}(a/q)\zeta_q^{ap}=\sum_{a=1}^{q-1}(ap^2/q)\zeta_q^{ap}=(p/q)\sum_{c=1}^{q-1}(c/q)\zeta_q^c=(p/q)z$, i. e. $(p/q)=1$. Cf. [5, Teil 2, § 55, Exercise 19].

[23]As an application we get the famous primality test for the Fermat numbers (cf. Footnote 29): $\mathsf{F}_t=2^{2^t}+1$, $t\ge1$, *is prime if and only if* $3^{(\mathsf{F}_t-1)/2}\equiv-1 \bmod \mathsf{F}_t$ (Pepin's Test). We have $\mathsf{F}_t\equiv(-1)^{2^t}+1\equiv2\equiv-1 \bmod 3$ and $\mathsf{F}_t\equiv1 \bmod 4$ for $t\ge1$ and hence, if F_t is prime, then $3^{(\mathsf{F}_t-1)/2}\equiv(3/\mathsf{F}_t)=(\mathsf{F}_t/3)=-1 \bmod \mathsf{F}_t$. Conversely, if $3^{(\mathsf{F}_t-1)/2}\equiv-1 \bmod \mathsf{F}_t$, then $\mathrm{ord}_{\mathsf{F}_t}3=\mathsf{F}_t-1$ and F_t is prime. Till today the only known Fermat primes are $\mathsf{F}_0,\mathsf{F}_1,\mathsf{F}_2,\mathsf{F}_3$ and F_4.

[24]By the way, this shows that $V=\pm q^{(q-1)/2}z$.

$\operatorname{sign} \mathrm{L}_p = (p/b)$ where (p/b) is the Jacobi symbol, cf. Footnote 20. Since $(bd-1)/2 \equiv (b+d-2)/2 \mod 2$ for arbitrary odd integers b,d we get $(p/bd) = (p/b)(p/d)$ for positive odd integers b,d not divisible by p and hence, quite generally,

$$\left(\frac{ac}{bd}\right) = \left(\frac{a}{b}\right)\left(\frac{c}{b}\right)\left(\frac{a}{d}\right)\left(\frac{c}{d}\right), \qquad (-1)^{\frac{(a-1)(b-1)}{4}}\left(\frac{b}{a}\right) = \left(\frac{a}{b}\right)$$

for arbitrary positive odd integers a,b,c,d *with* $\mathrm{GCD}(ac,bd) = 1$. This is the Jacobian version of the Quadratic Reciprocity Law which may also be deduced in a more formal way from the classical special case.

2.12 Diophantine Equations Nowadays, methods of modular arithmetic are of supreme importance in the study of Diophantine equations and have evolved into a far reaching theory. In this final subsection we give some examples which make use of the theory of quadratic residues.

2.12.1 Theorem *Let* p *be an odd prime number. The Diophantine equations* $x^2+y^2 = p$, $x^2-2y^2 = p$ *(or* $2y^2-x^2 = p$*) and* $x^2+2y^2 = p$ *have integer solutions if and only if* p *is congruent to* 1 *modulo* 4*, or congruent to* 1 *or* 7 *modulo* 8 *or to* 1 *or* 3 *modulo* 8*, respectively.*

To prove these results, consider the quadratic algebras $\mathbb{Z}[\sqrt{D}] = \mathbb{Z} \oplus \mathbb{Z}\sqrt{D} (\subseteq \mathbb{C})$ for $D = -1, 2, -2$, respectively, with their multiplicative norm functions $x+y\sqrt{D} \mapsto x^2 - y^2D = (x+y\sqrt{D})(x-y\sqrt{D}) = \mathrm{Det}(\mathrm{L}_{x+y\sqrt{D}})$ (where $\mathrm{L}_{x+y\sqrt{D}} : z \mapsto (x+y\sqrt{D})\,z$ denotes multiplication with $x+y\sqrt{D}$ in $\mathbb{Z}[\sqrt{D}]$), $x,y \in \mathbb{Z}$. All three algebras are Euclidean integral domains with respect to the absolute norm function $|x^2-y^2D|$. One checks this quite easily. In particular, they are principal ideal domains and factorial (i.e. UFDs). Furthermore, for an element $x+y\sqrt{D} \neq 0$ in $\mathbb{Z}[\sqrt{D}]$ the absolute norm $|x^2-y^2D|$ is the index of the principal ideal $\langle x+y\sqrt{D}\rangle = \mathbb{Z}[\sqrt{D}]\,(x+y\sqrt{D})$ in $\mathbb{Z}[\sqrt{D}]$. (More generally: For an injective group homomorphism $\varphi : \mathbb{Z}^m \to \mathbb{Z}^m$ given by an $m \times m$-matrix $\mathfrak{A} \in \mathrm{M}_m(\mathbb{Z})$ the absolute determinant $|\mathrm{Det}\,\varphi| = |\mathrm{Det}\,\mathfrak{A}| > 0$ is the index of the image $\mathrm{im}\,\varphi$ in $\mathbb{Z}^m$, i.e. the order of the cokernel $\mathrm{coker}\,\varphi = \mathbb{Z}^m/\mathrm{im}\,\varphi$. See [5, Teil 1, § 49, Corollary 49.8].) It follows: The Diophantine equation $|x^2-Dy^2| = n \in \mathbb{N}^*$, also called Pell's equation, has a solution $(x,y) \in \mathbb{Z}^2$ if and only if there exists a principal ideal of index n in $\mathbb{Z}[\sqrt{D}]$. Since any ideal of index n contains n, the ideals of index n in $\mathbb{Z}[\sqrt{D}]$ correspond bijectively to the ideals of index n in $\mathbb{Z}[\sqrt{D}]/\mathbb{Z}[\sqrt{D}]n = (\mathbb{Z}/\mathbb{Z}n)[\sqrt{D}] \cong (\mathbb{Z}/\mathbb{Z}n)[X]/(\mathbb{Z}/\mathbb{Z}n)[X](X^2-D)$. As a consequence we obtain: *If* $\mathbb{Z}[\sqrt{D}]$ *is a principal ideal domain, then, for a prime number* p*, Pell's equation* $|x^2-y^2D| = p$ *is solvable if and only if* D *is a square in* $\mathbb{F}_p = \mathbb{Z}/\mathbb{Z}p$, i. e. $(D/p) = 1$ or $p|D$. This proves all the claims made in the Theorem 2.12.1. For $D = 2$ one has to note that the solvability of $x^2-2y^2 = n$ and of $x^2-2y^2 = -n$ are equivalent conditions: (x,y) solves $x^2-2y^2 = n$ if and only if $(u,v) := (x+2y, x+y)$ solves $u^2-2v^2 = -n$. (The norm of $1+\sqrt{2}$ is -1.)

But, even more can be said: One easily calculates the Dedekind's ζ-function

$$\zeta_D(n) := \zeta_{\mathbb{Z}[\sqrt{D}]}(n),$$

i.e. the number of ideals of index $n \in \mathbb{N}^*$ in $\mathbb{Z}[\sqrt{D}]$. Since

$$\zeta_D(p_1^{\alpha_1} \cdots p_r^{\alpha_r}) = \zeta_D(p_1^{\alpha_1}) \cdots \zeta_D(p_r^{\alpha_r})$$

for pairwise distinct prime numbers $p_1, \ldots, p_r$, ζ_D is completely determined by the values $\zeta_D(p^\alpha) = 1$ if $p = 2$ or if p is prime with $p|D$, $\zeta_D(p^\alpha) = \alpha + 1$ if p is an odd prime number not dividing D with $(D/p) = 1$ and $\zeta_D(p^\alpha) = ((-1)^\alpha + 1)/2$ if p is an odd prime number not dividing D with $(D/p) = -1$. These formulae for the values of ζ_D are true for *all* square-free integers $D \not\equiv 1 \bmod 4$ because then $\mathbb{Z}[\sqrt{D}]$ is a so-called Dedekind domain, i. e. $\mathbb{Z}_{(p)}[\sqrt{D}] = \mathbb{Z}_{(p)} \oplus \mathbb{Z}_{(p)}\sqrt{D}$ is a principal ideal domain for every prime number p, where $\mathbb{Z}_{(p)} \subseteq \mathbb{Q}$ denotes the discrete valuation ring of those rational numbers whose (reduced) denominator is not divisible by p. $\mathbb{Z}_{(p)}[\sqrt{D}]$ has exactly

$$\begin{cases} \text{one (principal) maximal ideal of index } p \text{ if } p = 2 \text{ or } p|D, \\ \text{two (principal) maximal ideals of index } p \text{ each if } p \geq 3,\ p \not| D \text{ and } (D/p) = 1, \\ \text{one (principal) maximal ideal of index } p^2 \text{ if } p \geq 3,\ p \not| D \text{ and } (D/p) = -1. \end{cases}$$

This gives the values of the ζ-function ζ_D above.

Now, assume that $\mathbb{Z}[\sqrt{D}]$ even is a principal ideal domain. Then for $n \in \mathbb{N}^*$ the number of solutions of Pell's equation $|x^2 - y^2 D| = n \in \mathbb{N}^*$ is $e_D\, \zeta_D(n)$, where e_D is the order of the group of units in $\mathbb{Z}[\sqrt{D}]$, because two elements in $\mathbb{Z}[\sqrt{D}]$ generate the same ideal if and only if they differ multiplicatively by a unit. For $D > 0$ (D not a square) one always has $e_D = \infty$, i.e. $|x^2 - y^2 D| = n$ has no solution or infinitely many ones. For example, the units in $\mathbb{Z}[\sqrt{2}]$ are $\pm(1+\sqrt{2})^m$, $m \in \mathbb{Z}$. The units in $\mathbb{Z}[\sqrt{-1}]$ are $\pm 1, \pm\sqrt{-1}$, i.e. $e_{-1} = 4$, and for $D < -1$ the only units are ± 1, i.e. $e_D = 2$ for $D < -1$. So we get for $D = -1, -2$ and $n \in \mathbb{N}^*$ the following very precise results (recall that $v_p(n)$ denotes the exponent of the highest p-power which divides n, p prime):

2.12.2 Theorem (Two Squares Theorem of Euler-Fermat) *The number of solutions $(x,y) \in \mathbb{Z}^2$ of $x^2 + y^2 = n$ is equal to*

$$\begin{cases} 4 \cdot \prod\limits_{\substack{p|n, \\ p \equiv 1 \bmod 4}} (v_p(n) + 1), & \text{if } v_p(n) \text{ is even for all prime numbers } p \equiv 3 \bmod 4 \text{ dividing } n, \\ 0 & \text{else}. \end{cases}$$

2.12.3 Theorem *The number of solutions $(x,y) \in \mathbb{Z}^2$ of $x^2 + 2y^2 = n$ is equal to*

$$\begin{cases} 2 \cdot \prod\limits_{\substack{p|n, \\ p \equiv 1,3 \bmod 8}} (v_p(n) + 1), & \text{if } v_p(n) \text{is even for all prime numbers } p \equiv 5,\ 7 \bmod 8 \text{ dividing } n, \\ 0 & \text{else}. \end{cases}$$

For $D < -2$ the algebra $\mathbb{Z}[\sqrt{D}]$ is never a principal ideal domain because then $\mathbb{Z}[\sqrt{D}]$ contains (exactly) one ideal of index 2 (namely the ideal generated by 2 and $D + \sqrt{D}$), but no element $x + y\sqrt{D}$ with norm $x^2 + |D|y^2 = 2$. The first non-trivial algebra of this kind (with D square-free and $\not\equiv 1 \bmod 4$) is the famous example $\mathbb{Z}[\sqrt{-5}]$ of Dedekind.

For $D > 0$ the situation is different. For example, if $0 < D < 100$, then $\mathbb{Z}[\sqrt{D}]$ is a principal ideal domain if and only if $D \in \{2,3,6,7,11,14,19,22,23,31,38,43,46,47,59,62,67,71,83,86,94\}$. Cf. [5, Teil 2, § 59, Example 8] for a more thorough but still comparatively elementary investigation of the algebras $\mathbb{Z}[\sqrt{D}]$ and related rings.

§ 3 Special Shufflings

In this section we resume the discussion of shufflings which started in 1.1.

3.1 A shuffling machine shuffles a given pack $\mathbf{c} \in \mathfrak{P}_C$ of a set C of $n \in \mathbb{N}^*$ cards again and again with respect to a fixed permutation $\sigma \in \mathfrak{S}_n$, i. e. it produces successively the stacks

$$\mathbf{c}_0 = \mathbf{c},\ \mathbf{c}_1 = \sigma * \mathbf{c} = \mathbf{c}\sigma^{-1},\ \mathbf{c}_2 = \sigma * \mathbf{c}_1 = \sigma^2 * \mathbf{c} = \mathbf{c}\sigma^{-2},\ \ldots$$

$$\mathbf{c}_{i+1} = \sigma * \mathbf{c}_i = \sigma^{i+1} * \mathbf{c} = \mathbf{c}\sigma^{-(i+1)},\ \ldots\ ,$$

which is the orbit of the natural action on $\mathfrak{P}_C$ of the subgroup $\mathrm{H}(\sigma) \subseteq \mathfrak{S}_n$ generated by σ. Since the action of $\mathfrak{S}_n$ on $\mathfrak{P}_C$ is simply transitive and, in particular, free, the stacks $\mathbf{c}_i$ and $\mathbf{c}_j = \sigma^{j-i} * \mathbf{c}_i$ coincide if and only if $\sigma^{j-i} = \mathrm{id}_n$, i. e. if and only if $i \equiv j \bmod \mathrm{ord}\,\sigma$. This means *the sequence* $\mathbf{c}_0, \mathbf{c}_1, \mathbf{c}_2, \ldots$ *is* (*purely*) *periodic of period length* $\mathrm{ord}\,\sigma$.[25] In Example 1.7 we described a general method to compute the order of a permutation $\sigma \in \mathfrak{S}_n$ with the help of the cycle decomposition of σ. If one wishes that any card of C reaches in this sequence of stacks every position $1 \ldots, n$, i. e. that $\mathrm{H}(\sigma)$ operates transitively on $[1,n]$, then σ must be necessarily a cycle of length n, and, in particular, of order n. But, for $n > 4, n \neq 6$, there are permutations $\sigma \in \mathfrak{S}_n$ of order $> n$. By Example 1.7 the group $\mathfrak{S}_n$ contains an element of order $m \in \mathbb{N}^*$ if and only if $\sum_{p \in \mathbb{P},\, p|m} p^{v_p(m)} \leq n$.[26] It follows: The maximum of the orders of the elements of $\mathfrak{S}_n$ is

$$\mathrm{M}_n = \mathrm{Max}\,\{m \in \mathbb{N}^* \mid \sum_{p \in \mathbb{P},\, p|m} p^{v_p(m)} \leq n\}.$$

It seems to be rather cumbersome to compute the exact value of M_n for a given n. Obviously, one has $\mathrm{M}_1 \leq \mathrm{M}_2 \leq \mathrm{M}_3 \leq \cdots$. Compared to the exponent $\mathrm{Exp}\,\mathfrak{S}_n = \mathrm{LCM}\,(1,\ldots,n)$

[25] Let us fix the terminology for periodic sequences which is used here: For an arbitrary sequence $(x_i)_{i \in \mathbb{N}}$ of elements of a set X, a pair $(m_0, n) \in \mathbb{N} \times \mathbb{N}^*$ is called a pair of periodicity for (x_i) if $x_{i+n} = x_i$ for all $i \geq m_0$. In this case m_0 is called a pre-period length and n a period length of (x_i). If no such pair of periodicity for (x_i) exists, then (x_i) is called aperiodic, otherwise (x_i) is called periodic. One shows easily that, for a periodic sequence (x_i), there exists a *unique* pair of periodicity $(k_0, \ell) \in \mathbb{N} \times \mathbb{N}^*$ with the following property: Any pair of periodicity for (x_i) is of the form $(m_0, m\ell)$ with $m_0 \geq k_0$ and $m \in \mathbb{N}^*$. (The main point to show is the following: If $r, s \in \mathbb{N}^*$ are period lengths of (x_i), then $\mathrm{GCD}(r,s)$ is also a period length of (x_i).) One calls k_0 *the* pre-period length of (x_i) and ℓ *the* period length. The pair (k_0, ℓ) itself is called the (periodicity) type of (x_i). The (finite) sequence $(x_0, \ldots, x_{k_0-1})$ is *the* pre-period of (x_i) and $(x_{k_0}, \ldots, x_{k_0+\ell-1})$ *the* period. One writes shortly $(x_i)_{i \in \mathbb{N}} = (x_0, \ldots, x_{k_0-1}, \overline{x_{k_0}, \ldots, x_{k_0+\ell-1}})$. If $k_0 = 0$ then (x_i) is called purely periodic. The periodicity type of an aperiodic sequence is often denoted by $(\infty, 0)$. In particular, by definition, the period length of an aperiodic sequence is 0. – If x is an element of a group, the sequence $(x^i)_{i \in \mathbb{N}}$ of its powers has period length $\mathrm{ord}\,x$ and is purely periodic if $\mathrm{ord}\,x > 0$. For an element x of a monoid the periodicity type of the sequence $(x^i)_{i \in \mathbb{N}}$ characterizes the cyclic monoid generated by x up to isomorphism and any type in $\mathbb{N} \times \mathbb{N}^* \cup \{(\infty, 0)\}$ may occur. – For an integer $r \in \mathbb{N}^*$, the reader may compute the periodicity type of the sequence $(x_{ri})_{i \in \mathbb{N}}$ in terms of the periodicity type (k_0, ℓ) of $(x_i)_{i \in \mathbb{N}}$.

[26] In this formula $\mathbb{P} \subseteq \mathbb{N}^*$ denotes the set of all prime numbers and $m = \prod_{p \in \mathbb{P}} p^{v_p(m)}$ is the canonical prime decomposition of a positive integer m. For the proof of the next equality use the simple fact that, for arbitrary positive integers $m_1, \ldots, m_r \geq 2$, the inequality $m_1 + \cdots + m_r \leq m_1 \cdots m_r$ holds.

of the group $\mathfrak{S}_n$, cf. Footnote 15, the number M_n is rather small. Note that $B(n) := \mathrm{LCM}(1,\ldots,n) = e^{\psi(n)}$ where $\psi(n) = \sum_{m\le n}\Lambda(m)$, $n\in\mathbb{N}^*$, is the Chebyshev function ψ and Λ is the von Mangoldt function with $\Lambda(m) = \ln p$ if $m\in\mathbb{N}^*$ is a prime power $p^k \neq 1$ and $\Lambda(m) = 0$ else. Using the Prime Number Theorem $\psi(n)\sim n$ for $n\to\infty$ (i.e. $\lim_{n\to\infty}\psi(n)/n = 1$) one obtains the following estimates for $B(n) = \mathrm{Exp}\,\mathfrak{S}_n$: It is $\ln B(n) = \ln \mathrm{Exp}\,\mathfrak{S}_n \sim n$ for $n\to\infty$ and hence for arbitrary constants $C,D>1$ with $C<e<D$ one has $C^n < B(n) < D^n$ for almost all $n\in\mathbb{N}^*$. (For a discussion and a proof of the Prime Number Theorem cf. [7, Bd. 3, Beispiel 7.G.15], for example.)

3.1.1 Example We list the values of M_n for $n\le 20$:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
M_n	1	2	3	4	6	6	12	15	20	30	30	60	60	70	105	140	210	210	420	420

.

Furthermore, we mention that $\mathrm{M}_{49} = \mathrm{M}_{50} = \mathrm{M}_{51} = \mathrm{M}_{52} = 2^2\cdot 3^2\cdot 5\cdot 7\cdot 11\cdot 13 = 180\,180$ and $\mathrm{M}_{53} = 2\cdot \mathrm{M}_{52} = 360\,360$. Compare this with $\mathrm{Exp}\,\mathfrak{S}_{53} = 164\,249\,358,725\,037\,825\,439\,200 \approx 164\cdot 10^{21}$.

3.2 Discrete Logarithm Problem for $\mathfrak{S}_n$ Another natural question is the following: Given a stack $\mathbf{d}\in\mathfrak{P}_C$, does $\mathbf{d}$ occur in the sequence $\mathbf{c} = \mathbf{c}_0,\mathbf{c}_1,\mathbf{c}_2,\ldots$, i.e. can $\mathbf{d}$ be obtained from $\mathbf{c}$ by iterated shuffling with the given permutation $\sigma\in\mathfrak{S}_n$, or equivalently, is $\mathbf{d} = \sigma^x * \mathbf{c} = \mathbf{c}\sigma^{-x}$, i.e. $\tau := \mathbf{d}^{-1}\mathbf{c} = \sigma^x$ for some $x\in\mathbb{N}$. Thus, for the answer we need to solve the discrete logarithm problem (DLP) belonging to the data $(\mathfrak{S}_n;\sigma,\tau)$, which we already mentioned in a more general setting in Example 2.8. Usually the cycle decomposition of σ yields a representation of $\mathrm{ord}\,\sigma$ as a product of small positive integers. Therefore one may expect, after the discussions in Example 2.8, that for moderate $n = \#C$ the problem is rather easy to solve. However, here we give a direct method. We compute the cycle decomposition of σ (cf. Example 1.7):

$$\sigma = \sigma_1\cdots\sigma_r = \langle a_0^{(1)},\ldots,a_{m_1-1}^{(1)}\rangle\cdots\langle a_0^{(r)},\ldots,a_{m_r-1}^{(r)}\rangle .$$

The orbits of σ are invariant under σ and hence, if the DLP $(\mathfrak{S}_n;\sigma,\tau)$ has a solution, the orbits have to be invariant under τ, too. If this is the case, then

$$\tau(a_0^{(\rho)}) = a_{x_\rho}^{(\rho)} \quad \text{for some } x_\rho \text{ with } \quad 0\le x_\rho < m_\rho\,,\ \rho = 1,\ldots,r.$$

Now, if there exists an x with $\tau = \sigma^x$, then necessarily

$$x\equiv x_\rho \mod m_\rho\,,\quad \rho = 1,\ldots,r,$$

because $\sigma^x(a_0^{(\rho)}) = a_{x_\rho}^{(\rho)}$ if and only if $x\equiv x_\rho \mod m_\rho$. By the Generalized Chinese Remainder Theorem 2.10.1, this system of simultaneous congruences has a solution if and only if the solvability conditions

$$x_\lambda \equiv x_\mu \mod \mathrm{GCD}(m_\lambda,m_\mu),\quad 1\le\lambda<\mu\le r,$$

hold. If these necessary conditions hold, then there exists a solution x with $0\le x < \mathrm{ord}\,\sigma = \mathrm{LCM}(m_1,\ldots,m_r)$. It is uniquely determined and may be computed rather easily,

cf. 2.10.2. Hence, if the DLP $(\mathfrak{S}_n;\sigma,\tau)$ has a solution, then $\log_\sigma \tau = x$. Now, one checks the validity of the equation

$$\sigma^x = \sigma_1^x \cdots \sigma_r^x = \sigma_1^{x_1} \cdots \sigma_r^{x_r} = \tau$$

(which is possible without knowing x explicitly).

3.3 Faro Shufflings Now we discuss some special shuffling methods which are well-known to card players. The Faro shuffling[27] is performed by cutting the given pack into two equal piles (as far as possible) and then by taking alternately one card from each pile to form the new pack. Therefore, if the number n of cards is even, $n = 2m$, then there are two possibilities which are described by the following permutations $\sigma_{1A}, \sigma_{1B} \in \mathfrak{S}_{2m}$ (according to the convention (1) of Subsection 1.1):

$$\sigma_{1A} := \begin{pmatrix} 1 & 2 & 3 & \dots & m & m+1 & m+2 & \dots & 2m-1 & 2m \\ 1 & 3 & 5 & \dots & 2m-1 & 2 & 4 & \dots & 2(m-1) & 2m \end{pmatrix},$$

$$\sigma_{1B} := \begin{pmatrix} 1 & 2 & 3 & \dots & m & m+1 & m+2 & \dots & 2m-1 & 2m \\ 2 & 4 & 6 & \dots & 2m & 1 & 3 & \dots & 2m-3 & 2m-1 \end{pmatrix}.$$

If n is odd, $n = 2m+1$, there are four possibilities described by the following permutations $\sigma_{2A}, \sigma_{2B}; \sigma_{3A}, \sigma_{3B} \in \mathfrak{S}_{2m+1}$ (in this case one pile contains m and the other $m+1$ cards):

$$\sigma_{2A} := \begin{pmatrix} 1 & 2 & 3 & \dots & m & m+1 & m+2 & \dots & 2m & 2m+1 \\ 1 & 3 & 5 & \dots & 2m-1 & 2 & 4 & \dots & 2m & 2m+1 \end{pmatrix},$$

$$\sigma_{2B} := \begin{pmatrix} 1 & 2 & 3 & \dots & m & m+1 & m+2 & \dots & 2m & 2m+1 \\ 2 & 4 & 6 & \dots & 2m & 1 & 3 & \dots & 2m-1 & 2m+1 \end{pmatrix},$$

$$\sigma_{3A} := \begin{pmatrix} 1 & 2 & 3 & \dots & m+1 & m+2 & m+3 & \dots & 2m & 2m+1 \\ 1 & 3 & 5 & \dots & 2m+1 & 2 & 4 & \dots & 2(m-1) & 2m \end{pmatrix},$$

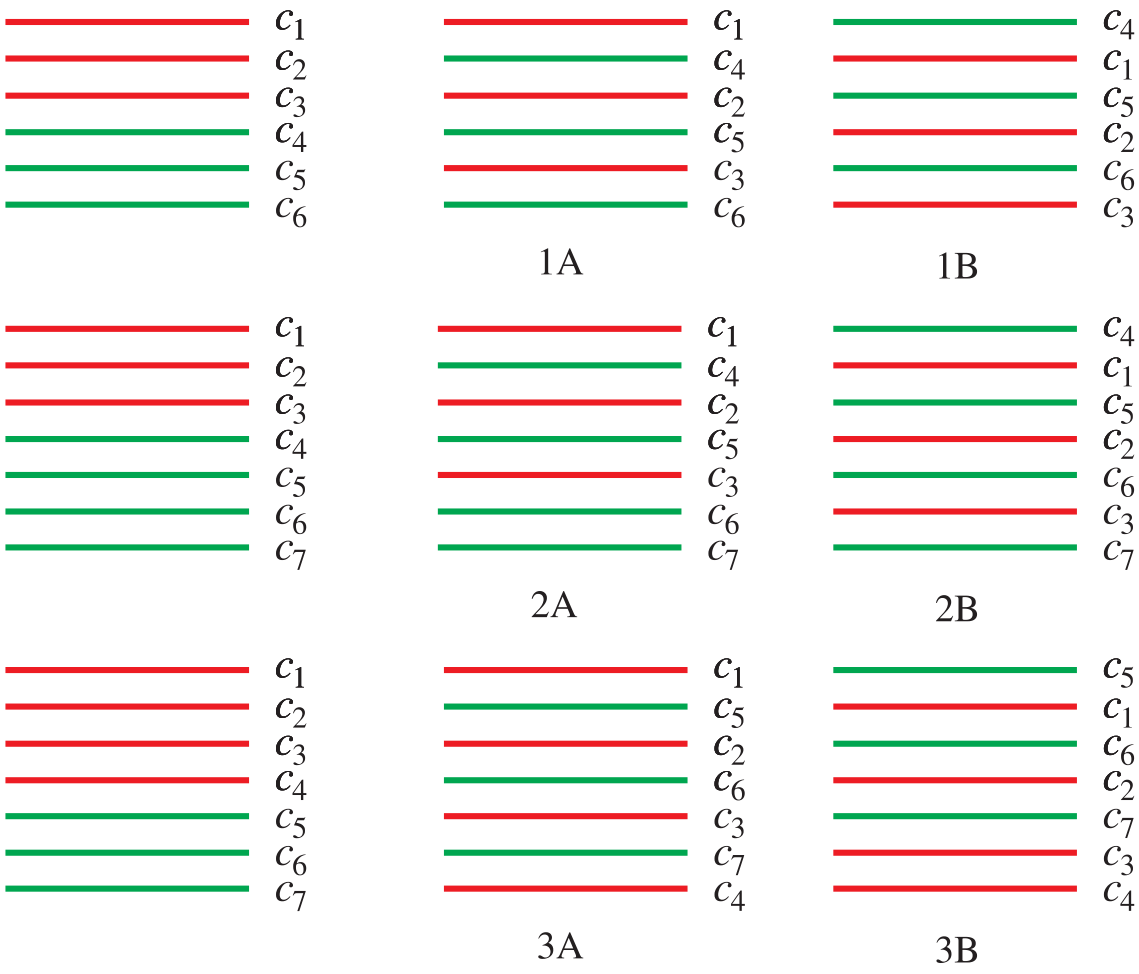
$$\sigma_{3B} := \begin{pmatrix} 1 & 2 & 3 & \dots & m+1 & m+2 & m+3 & \dots & 2m & 2m+1 \\ 2 & 4 & 6 & \dots & 2m+1 & 1 & 3 & \dots & 2m-3 & 2m-1 \end{pmatrix}.$$

The Faro shufflings which leave the original top card at the top are known as out-shuffling s. These are the cases $1A, 2A, 3A$. The Faro shufflings which move the original top card to the second place are known as in-shuffling s. These are the cases $1B, 2B, 3B$. Faro shufflings are also described in [1] and [2].

3.3.1 Example We illustrate the Faro shuffling by giving the pictures of the resulting packs after shuffling the pack $\mathbf{c} = (c_1, \dots, c_n)$ for $n = 6$ and $n = 7$:

---

[27] The name goes back to the Faro (or Pharo or Pharaoh) game. Many sources say the game of Faro originated in France in the early 18th century (about 1713) as a revised form of the popular British pub game basset. Basset was outlawed in France by King Louis XIV in 1691, and Faro was developed by European gamblers as an alternative. Although both Faro and Basset were forbidden in France, these games continued to be popular in England. In 19th century Faro was the most commonly played card game in the Old West of America.



3.3.2 Remark The permutations of the Faro shufflings belong to the general shuffle permutations of $\mathfrak{S}_n$. A shuffle permutation $\sigma_R \in \mathfrak{S}_n$ is characterized by a subset $R \subseteq [1,n]$. If $R = \{i_1, i_2, \ldots, i_r\}$ with $i_1 < i_2 < \cdots < i_r$, then shuffling with permutation σ_R puts the upper r cards of the given pack to the positions $i_1, i_2, \ldots, i_r$, and the remaining $s := n - r$ cards to the positions $j_1, j_2, \ldots, j_s$, respectively, where $\{j_1, j_2, \ldots, j_s\} = [1,n] \setminus R$ and $j_1 < j_2 < \cdots < j_s$, i. e.

$$\sigma_R = \begin{pmatrix} 1 & 2 & \ldots & r & r+1 & \ldots & r+s \\ i_1 & i_2 & \ldots & i_r & j_1 & \ldots & j_s \end{pmatrix}.$$

Note that the definition of σ_R uses not only the subset R, but also the canonical order on the set $[1,n]$. For a fixed r, $0 \le r \le n$, the $\binom{n}{r}$ shuffle permutations σ_R with $\#R = r$ form a canonical system of representatives for the *left* cosets of the subgroup $\mathfrak{S}_{r,n-r} := \mathfrak{S}([1,r]) \times \mathfrak{S}([r+1,n])$ in $\mathfrak{S}_n$: For $\sigma \in \mathfrak{S}_n$, the equality $\sigma\mathfrak{S}_{r,n-r} = \sigma_R\mathfrak{S}_{r,n-r}$ holds for the unique subset $R := \sigma([1,r]) \subseteq [1,n]$ of cardinality r. (The inverses σ_R^{-1}, $\#R = r$, form a system of representatives for the *right* cosets of $\mathfrak{S}_{r,n-r}$ in $\mathfrak{S}_n$.)

In order to understand the permutations σ of Faro shufflings, we try to interpret them as permutations of sets X with additional structures, i. e. instead of a given σ we consider the conjugated permutation $\varphi\sigma\varphi^{-1}$ for some appropriate bijection $\varphi : [1,n] \to X$, cf. Proposition 1.8.

To handle σ_{1A}, we omit the fixed point $2m$ and consider σ_{1A} as an element of $\mathfrak{S}_{2m-1}$. Now, interpreting the integers $1,\dots,2m-1$ as the elements of the ring $\mathbb{Z}_{2m-1}$, σ_{1A} is obviously the affine transformation $\mathbb{Z}_{2m-1}\to\mathbb{Z}_{2m-1}$, $x\mapsto 2x-1$. Since it has the fixed point 1 it is conjugated to the homothecy $\vartheta_2 : x\mapsto 2x$ and hence

$$\operatorname{ord}\sigma_{1A} = \operatorname{ord}\vartheta_2 = \operatorname{ord}_{2m-1} 2\,.$$

The permutation σ_{1B} can directly be interpreted as the homothecy $\vartheta_2 : \mathbb{Z}_{2m+1}\to\mathbb{Z}_{2m+1}$ and hence

$$\operatorname{ord}\sigma_{1B} = \operatorname{ord}\vartheta_2 = \operatorname{ord}_{2m+1} 2\,.$$

3.3.3 Example Bridge[28] is a well-known game played with the standard set of $n=2m=52$ cards. For the out-shuffling 1A, we need to compute $\operatorname{ord}_{51} 2 = \operatorname{LCM}(\operatorname{ord}_3 2, \operatorname{ord}_{17} 2) = \operatorname{ord}_{17} 2 = 8$, since $\operatorname{ord}_3 2 = 2$, $2^4\equiv -1 \bmod 17$, $2^8\equiv 1 \bmod 17$.[29] – For the in-shuffling 1B, we need to compute $\operatorname{ord}_{53} 2$ which divides $\operatorname{ord}\mathbb{Z}_{53}^\times = 53-1 = 52 = 2^2\cdot 13$. Since $2^{2^2}\not\equiv 1 \bmod 53$ and, by 2.11.5 $2^{2\cdot 13}\equiv (2/53) = -1\not\equiv 1 \bmod 53$, we have $\operatorname{ord}_{53} 2 = 2^2\cdot 13 = 52$, cf. Example 2.8.

The cases 2A and 2B are treated similarly 1A and 1B, respectively, with the result

$$\operatorname{ord}\sigma_{2A} = \operatorname{ord}_{2m-1} 2 \quad\text{and}\quad \operatorname{ord}\sigma_{2B} = \operatorname{ord}_{2m+1} 2\,.$$

Further, obviously, $\operatorname{ord}\sigma_{3A} = \operatorname{ord}\sigma_{2B}$, and hence

$$\operatorname{ord}\sigma_{3A} = \operatorname{ord}_{2m+1} 2\,.$$

3.3.4 Example The card game Old Maid[30] is played with a set of $n = 33 = 2\cdot 16+1$ picture cards. In this case $\operatorname{ord}\sigma_{2A} = \operatorname{ord}_{31} 2 = 5$, since $2^5 = 32\equiv 1 \bmod 31$. $\operatorname{ord}\sigma_{2B} = \operatorname{ord}\sigma_{3A} = \operatorname{ord}_{33} 2 = \operatorname{LCM}(\operatorname{ord}_3 2, \operatorname{ord}_{11} 2) = \operatorname{LCM}(2,10) = 10$, since $2^2\not\equiv 1 \bmod 11$ and $2^5\equiv -1\not\equiv 1 \bmod 11$.

The case 3B is the most difficult one. Instead of σ_{3B} we consider the permutation

$$\sigma' := \begin{pmatrix} 0 & 1 & 2 & \dots & m & m+1 & m+2 & \dots & 2m & 2m+1 & 2m+2 \\ 0 & 2 & 4 & \dots & 2m & 2m+1 & 1 & \dots & 2m-3 & 2m-1 & 2m+2 \end{pmatrix}$$

by adding two fixed points 0 and $2m+2$. Obviously,

$$\operatorname{ord}\sigma_{3B} = \operatorname{ord}\sigma'\,.$$

---

[28] No one knows precisely where the name "Bridge" for the card game comes from, although it is fairly certain that it has nothing to do with other meanings of the word "bridge". One proposal for the etymology of the word "Bridge" for the card game is the following: In the 19th century in eastern countries a card game was popular which was called "Whist" in Russia. This game was also known as Biritch or Britch. Both these words sound Russian although neither of them seems to be Russian. Anyway, once the British took up the game (and changed the rules), "Britch" became "Bridge" through folk etymology.

[29] More generally, $\operatorname{ord}_{2^{2^t}+1} 2 = 2^{t+1}$, since $2^{2^t}\equiv -1\not\equiv 1 \bmod 2^{2^t}+1$ and $2^{2^{t+1}}\equiv 1 \bmod 2^{2^t}+1$. The numbers $\mathsf{F}_t := 2^{2^t}+1$, $t\geq 0$, are called Fermat numbers. For any prime factor p of F_t we also have $\operatorname{ord}_p 2 = 2^{t+1}$ and, in particular, $p\equiv 1 \bmod 2^{t+1}$, hence, by Proposition 2.11.5, $(2/p) = 1$ for $t>1$, which implies $2^{t+1}|(p-1)/2$ or $p\equiv 1 \bmod 2^{t+2}$. For example, for $t=5$, the first primes $\equiv 1 \bmod 2^7$ are $\mathsf{F}_3 = 2\cdot 2^7+1$ and $641 = 5\cdot 2^7+1$, and, indeed, $641\,|\,\mathsf{F}_5$ because $641 = 5^4+2^4$, i. e. $2^4\equiv -5^4 \bmod 641$, hence $2^{32}\equiv -5^4\cdot 2^{28}\equiv -(5\cdot 2^7)^4\equiv -1 \bmod 641$. The next Fermat number F_6 is divisible by $1071\cdot 2^8+1$.

[30] Old Maid is a traditional children card game in most English-speaking countries. The set of cards consists of 16 pairs and a single card called Old Maid, usually featuring an old woman or a spinster. By picking cards in turn from each other's hands, the players try to discard couples and to avoid being left with the Old Maid, whose holder at the end is the loser. The same game is even more popular in German-speaking countries and those related to German culture, where the game is known as Schwarzer Peter (German), Cerný Petr (Czech), Cierny Peter (Slovak), Fekete Péter (Hungarian), Uomo Nero (Italian), ... In these varieties the single card of the game is often featuring a chimney-sweeper (Black Peter) who is considered as a mascot.

Now, in σ', we interchange the images of $m+1$ and $2m+2$, and get the permutation $\sigma'' = \langle 2m+1, 2m+2\rangle \cdot \sigma'$ with

$$\sigma'' = \begin{pmatrix} 0 & 1 & 2 & \dots & m & m+1 & m+2 & \dots & 2m & 2m+1 & 2m+2 \\ 0 & 2 & 4 & \dots & 2m & 2m+2 & 1 & \dots & 2m-3 & 2m-1 & 2m+1 \end{pmatrix}$$

which is visibly multiplication by 2 in the ring $\mathbb{Z}_{2m+3}$ and so $\operatorname{ord} \sigma'' = \operatorname{ord}_{2m+3} 2$.

Let $\ell_1, \dots, \ell_r$ be the cardinalities of the orbits of σ' which are not singletons and let ℓ_r be the cardinality of the orbit $\mathrm{H}(\sigma') \cdot (2m+1)$ (which is not a singleton). Then $\ell_1, \dots, \ell_{r-1}, \ell_r + 1$ are the cardinalities of the orbits of σ'' which are not singletons, and $\ell_r + 1$ is the cardinality of the orbit $\mathrm{H}(\sigma'') \cdot (2m+2) = \mathrm{H}(2) \cdot (-1)$, since $2m+2 = -1$ in $\mathbb{Z}_{2m+3}$. ($\mathrm{H}(2)$ is the subgroup generated by 2 in $\mathbb{Z}^\times_{2m+2}$.) We get $\ell_r + 1 = \#\mathrm{H}(2) = \operatorname{ord}_{2m+3} 2$. Now, we distinguish the following two cases:

a) The element 2 does not generate the prime residue class group $\mathbb{Z}^\times_{2m+3}$, i. e. there exists an element $a \in \mathbb{Z}^\times_{2m+3} \setminus \mathrm{H}(2) \cdot (-1)$, the orbit $\mathrm{H}(\sigma'') \cdot a = \mathrm{H}(2) \cdot a$ of which has also cardinality $\#\mathrm{H}(2) = \operatorname{ord}_{2m+3} 2$. This is one of the numbers $\ell_1, \dots, \ell_{r-1}$, and all the others are divisors of $\#\mathrm{H}(2)$ by Theorem 1.3. Therefore it follows that $\operatorname{ord} \sigma'$ $(= \operatorname{ord} \sigma_{3\mathrm{B}})$ is equal to

$$\begin{aligned} \operatorname{LCM}(\ell_1, \dots, \ell_{r-1}, \ell_r) &= \operatorname{LCM}(\operatorname{ord}_{2m+3} 2, (\operatorname{ord}_{2m+3} 2) - 1) \\ &= \operatorname{ord}_{2m+3} 2 \cdot (\operatorname{ord}_{2m+3} 2 - 1). \end{aligned}$$

b) The element 2 generates the prime residue class group $\mathbb{Z}^\times_{2m+3}$. Then $\mathbb{Z}^\times_{2m+3} = \mathrm{H}(2)$ is a cyclic group and, by Theorem 2.7, $2m+3 = p^\alpha$ with a prime number $p \geq 3$. If $\alpha = 1$, then $r = 1$ and hence

$$\operatorname{ord} \sigma_{3\mathrm{B}} = \operatorname{ord} \sigma' = \operatorname{ord} \sigma'' - 1 = \operatorname{ord}_p 2 - 1 = p - 2 = 2m+1.$$

i. e. $\sigma_{3\mathrm{B}}$ is a cycle of length $2m+1$. Now, let $\alpha \geq 2$. The orbits of σ'' are the orbits of the canonical operation $\mathbb{Z}^\times_{p^\alpha} \times \mathbb{Z}_{p^\alpha} \to \mathbb{Z}_{p^\alpha}$ of $\mathbb{Z}^\times_{p^\alpha} = \mathrm{H}(2)$ on $\mathbb{Z}_{p^\alpha}$ by multiplication which we discussed already in Example 2.9. Each orbit is formed by the elements of a fixed order p^β, $0 \leq \beta \leq \alpha$. Hence, the sequence $\ell_1, \dots, \ell_{r-1}, \ell_r$ coincides (up to permutation of the elements $\ell_1, \dots, \ell_{r-1}$) with the sequence $(p-1), p(p-1), \dots, p^{\alpha-2}(p-1), p^{\alpha-1}(p-1)$. It follows

$$\operatorname{ord} \sigma_{3\mathrm{B}} = \operatorname{ord} \sigma' = \operatorname{LCM}(\ell_1, \dots, \ell_{r-1}, \ell_r) = p^{\alpha-2}\big(p-1\big)\big(p^{\alpha-1}(p-1) - 1\big).$$

Altogether, we have completed the proof of the following theorem:

3.4 Theorem *The Faro shuffling permutations of a pack of $n \in \mathbb{N}^*$ cards as described in the beginning of* Subsection 3.3 *have the following orders*:

(1) *If $n = 2m$ is even, then* $\operatorname{ord} \sigma_{1\mathrm{A}} = \operatorname{ord}_{2m-1} 2$ *and* $\operatorname{ord} \sigma_{1\mathrm{B}} = \operatorname{ord}_{2m+1} 2$.
(2) *If $n = 2m+1$ is odd, then* $\operatorname{ord} \sigma_{2\mathrm{A}} = \operatorname{ord}_{2m-1} 2$, $\operatorname{ord} \sigma_{2\mathrm{B}} = \operatorname{ord} \sigma_{3\mathrm{A}} = \operatorname{ord}_{2m+1} 2$ *and*

$$\operatorname{ord} \sigma_{3\mathrm{B}} = \begin{cases} \operatorname{ord}_{2m+3} 2 \cdot (\operatorname{ord}_{2m+3} 2 - 1), & \text{if } \mathrm{H}(2) \neq \mathbb{Z}^\times_{2m+3}, \\ 2m+1 = p-2, & \text{if } 2m+3 =: p \text{ is prime and} \\ & \quad \mathrm{H}(2) = \mathbb{Z}^\times_p, \\ p^{\alpha-2}\big(p-1\big)\big(p^{\alpha-1}(p-1) - 1\big), & \text{if } 2m+3 =: p^\alpha,\ p \text{ prime},\ \alpha \geq 2, \\ & \quad \text{and } \mathrm{H}(2) = \mathbb{Z}^\times_{p^\alpha}. \end{cases}$$

(Note that $\mathrm{H}(2)$ is the subgroup generated by 2.)

3.4.1 Example Continuing Example 3.3.4, to compute $\operatorname{ord}\sigma_{3\mathrm{B}}$ for $n=2m+1=33$, we have to apply the first case of the formula for $\operatorname{ord}\sigma_{3\mathrm{B}}$ in Theorem 3.4(2), since 2 is not a generator of the (non-cyclic) group $\mathbb{Z}_{35}^{\times}$. Hence $\operatorname{ord}\sigma_{3\mathrm{B}}=\operatorname{ord}_{35}2\cdot(\operatorname{ord}_{35}2-1)=12\cdot 11=132$ because of $\operatorname{ord}_{35}2=\mathrm{LCM}(\operatorname{ord}_5 2,\operatorname{ord}_7 2)=\mathrm{LCM}(4,3)=12$.

3.4.2 Example Theorem 3.4 implies in particular: A Faro permutation $\sigma\in\mathfrak{S}_n$ operates transitively (i. e. σ is a cycle of length n) only in the following cases:

(1) $n=2m$ is even, $p:=n+1=2m+1$ is a prime number with $\mathrm{H}(2)=\mathbb{Z}_p^{\times}$ and $\sigma=\sigma_{1\mathrm{B}}\in\mathfrak{S}_n$ belongs to the in-shuffling case 1B.

(2) $n=2m+1$ is odd, $p:=n+2=2m+3$ is a prime number with $\mathrm{H}(2)=\mathbb{Z}_p^{\times}$ and $\sigma=\sigma_{3\mathrm{B}}\in\mathfrak{S}_n$ belongs to the in-shuffling case 3B.

If 2 is, as in (1) and (2) above, a primitive prime residue modulo the odd prime number p, then, by the theory of quadratic residues (see Proposition 2.11.5), $p\equiv 3 \bmod 8$ or $p\equiv 5 \bmod 8$. But, the only odd prime numbers <200 with $p\equiv 3$ or $5 \bmod 8$ for which 2 is *not* a primitive prime residue modulo p are $p=43$, 109 and 157 (with $\operatorname{ord}_{43}2=14$, $\operatorname{ord}_{109}2=36$, $\operatorname{ord}_{157}2=52$). Hence, the only Faro permutations in $\mathfrak{S}_n$, $1\le n\le 200$, which operate transitively are $\sigma_{1\mathrm{B}}$ for $n=2m=p-1$ and $\sigma_{3\mathrm{B}}$ for $n=2m+1=p-2$, where p is a prime number with $p\equiv 3$ or $5 \bmod 8$ and $p\neq 43, 109, 157$.

3.4.3 Example The orders of the six Faro shufflings illustrated in Example 3.3.1 are successively $\operatorname{ord}_5 2=4$, $\operatorname{ord}_7 2=3$, $\operatorname{ord}_5 2=4$, $\operatorname{ord}_7 2=3$, $\operatorname{ord}_7 2=3$, $(3-1)(3(3-1)-1)=10$. – In connection with the computation of $\operatorname{ord}\sigma_{3\mathrm{B}}$ in the case that $n+2=2m+3=p^{\alpha}$ is an odd prime power, the problem arises how to decide whether 2 is a primitive prime residue modulo p^{α}. Now, if 2 is a primitive prime residue modulo p^{α} for $\alpha>0$, then 2 is a primitive prime residue modulo p^{β} for every β, $1\le\beta\le\alpha$, in particular $p\equiv 3 \bmod 8$ or $p\equiv 5 \bmod 8$, see the previous example for the case $\alpha=1$. For $\alpha\ge 2$, the following Proposition 3.5 may be useful.

3.5 Proposition *Let p be an odd prime number and let $\alpha\ge 2$. The following conditions are equivalent*:

(i) *2 is a primitive prime residue modulo p^{α}.*

(ii) *2 is a primitive prime residue modulo p^2.*

(iii) *2 is a primitive prime residue modulo p and $2^{p-1}\not\equiv 1 \bmod p^2$.*

The proof is left to the reader, cf. the proof of Theorem 2.5. By the way, in Proposition 3.5 the number 2 in the statements (i), (ii), (iii) can be replaced (simultaneously) by any integer not divisible by p. No odd prime number p is known for which 2 is a primitive prime residue modulo p, but not a primitive prime residue modulo p^{α}, $\alpha\ge 2$.[31]

3.6 Monge Shufflings The Monge shuffling[32] takes cards from a given pack

[31] Indeed, the only known odd prime numbers p with $2^{p-1}\equiv 1 \bmod p^2$ (which are called Wieferich primes) are 1093 ($\equiv 5 \bmod 8$, found by Meissner in 1913) and 3511 ($\equiv 7 \bmod 8$, found by Beeger in 1922), but, for these two prime numbers, 2 is not a primitive prime residue modulo p. Any other Wieferich prime must be $>1.25\cdot 10^{15}$ (as checked with a computer by Knauer/Richstein in 2003) or even $>2.5\cdot 10^{15}$ (cf. Ribenboim, R.: Meine Zahlen, meine Freunde. Springer, Berlin/Heidelberg 2009, p. 258). It is not known whether there are infinitely many Wieferich primes or whether there are infinitely many non-Wieferich primes.

[32] Named after Gaspard Monge (1746-1818) who investigated this shuffling in 1773, cf. G. Monge: Réflexions sur un tour des cartes, Mém. math. phys. présentés à l'Académie des Sciences, Paris (1773), 390-412.

$\mathbf{c} = (c_1, \ldots, c_n)$, alternately a card from the top and from the bottom, to form the new pack

$$\mathbf{c}_{\mathrm{A}} := \begin{cases} (c_1, c_{2m}, c_2, c_{2m-1}, \ldots, c_m, c_{m+1}), & \text{if } n = 2m \text{ is even}, \\ (c_1, c_{2m+1}, c_2, c_{2m}, \ldots, c_m, c_{m+2}, c_{m+1}), & \text{if } n = 2m+1 \text{ is odd}. \end{cases}$$

3.6.1 Example If $n = 6$ and $\mathbf{c} = (1,2,3,4,5,6)$, then $\mathbf{c}_{\mathrm{A}} = (1,6,2,5,3,4)$, and, if $n = 7$ and $\mathbf{c} = (1,2,3,4,5,6,7)$, then $\mathbf{c}_{\mathrm{A}} = (1,7,2,6,3,5,4)$.

The Monge shuffling is performed with the Monge permutation

$$\sigma_{\mathrm{MA}} := \begin{cases} \begin{pmatrix} 1 & 2 & \ldots & m & m+1 & m+2 & \ldots & 2m-1 & 2m \\ 1 & 3 & \ldots & 2m-1 & 2m & 2(m-1) & \ldots & 4 & 2 \end{pmatrix}, & \text{if } n = 2m \text{ is even}, \\ \begin{pmatrix} 1 & 2 & \ldots & m & m+1 & m+2 & \ldots & 2m & 2m+1 \\ 1 & 3 & \ldots & 2m-1 & 2m+1 & 2m & \ldots & 4 & 2 \end{pmatrix}, & \text{if } n = 2m+1 \text{ is odd}. \end{cases}$$

(Remember that we use the convention (1) of Subsection 1.1.) Analogously, from a given pack $\mathbf{c} = (c_1, \ldots, c_n)$ one can take alternately a card first from the bottom and then from the top to form the new pack

$$\mathbf{c}_{\mathrm{B}} = \begin{cases} (c_{2m}, c_1, c_{2m-1}, c_2, \ldots, c_{m+1}, c_m), & \text{if } n = 2m \text{ is even}, \\ (c_{2m+1}, c_1, c_{2m}, c_2, \ldots, c_{m+2}, c_m, c_{m+1}), & \text{if } n = 2m+1 \text{ is odd}. \end{cases}$$

3.6.2 Example If $n = 6$ and $\mathbf{c} = (1,2,3,4,5,6)$, then $\mathbf{c}_{\mathrm{B}} = (6,1,5,2,4,3)$, and, if $n = 7$ and $\mathbf{c} = (1,2,3,4,5,6,7)$, then $\mathbf{c}_{\mathrm{B}} = (7,1,6,2,5,3,4)$.

The corresponding Monge permutation is now

$$\sigma_{\mathrm{MB}} := \begin{cases} \begin{pmatrix} 1 & 2 & \ldots & m & m+1 & m+2 & \ldots & 2m-1 & 2m \\ 2 & 4 & \ldots & 2m & 2m-1 & 2m-3 & \ldots & 3 & 1 \end{pmatrix}, & \text{if } n = 2m \text{ is even}, \\ \begin{pmatrix} 1 & 2 & \ldots & m & m+1 & m+2 & \ldots & 2m & 2m+1 \\ 2 & 4 & \ldots & 2m & 2m+1 & 2m-1 & \ldots & 3 & 1 \end{pmatrix}, & \text{if } n = 2m+1 \text{ is odd}. \end{cases}$$

The permutation σ_{MA} describes a Monge out-shuffling and the permutation σ_{MB} a Monge in-shuffling. The Monge shufflings are also mentioned in [1].

We prove the following theorem:

3.7 Theorem *The Monge shufflings of a pack of $n \in \mathbb{N}^*$ cards have the following orders*:

(1) *If $n = 2m$ is even, then*

$$\operatorname{ord} \sigma_{\mathrm{MA}} = \begin{cases} \operatorname{ord}_{4m-1} 2, & \text{if } -1 \notin \mathrm{H}(2) \subseteq \mathbb{Z}_{4m-1}^{\times}, \\ \frac{1}{2} \operatorname{ord}_{4m-1} 2, & \text{if } -1 \in \mathrm{H}(2) \subseteq \mathbb{Z}_{4m-1}^{\times}, \end{cases}$$

$$\operatorname{ord}\sigma_{\mathrm{MB}} = \begin{cases} \operatorname{ord}_{4m+1} 2, & \text{if } -1 \notin \mathrm{H}(2) \subseteq \mathbb{Z}_{4m+1}^{\times}, \\ \frac{1}{2}\operatorname{ord}_{4m+1} 2, & \text{if } -1 \in \mathrm{H}(2) \subseteq \mathbb{Z}_{4m+1}^{\times}. \end{cases}$$

(2) *If* $n = 2m+1$ *is odd and* ≥ 3*, then*

$$\operatorname{ord}\sigma_{\mathrm{MA}} = \begin{cases} \operatorname{ord}_{4m+1} 2, & \text{if } -1 \notin \mathrm{H}(2) \subseteq \mathbb{Z}_{4m+1}^{\times}, \\ \frac{1}{2}\operatorname{ord}_{4m+1} 2, & \text{if } -1 \in \mathrm{H}(2) \subseteq \mathbb{Z}_{4m+1}^{\times}, \end{cases}$$

$$\operatorname{ord}\sigma_{\mathrm{MB}} = \begin{cases} \operatorname{ord}_{4m+3} 2, & \text{if } -1 \notin \mathrm{H}(2) \subseteq \mathbb{Z}_{4m+3}^{\times}, \\ \frac{1}{2}\operatorname{ord}_{4m+3} 2, & \text{if } -1 \in \mathrm{H}(2) \subseteq \mathbb{Z}_{4m+3}^{\times}. \end{cases}$$

To prove 3.7 we start with a general remark: Let $k \in \mathbb{N}^*$ be an odd positive integer ≥ 3. Then the subgroup $\{\pm 1\} \subseteq \mathbb{Z}_k^{\times}$ of order 2 operates canonically on $\mathbb{Z}_k$. The set of orbits $\overline{X}_k := \mathbb{Z}_k \backslash \{\pm 1\}$ contains the singleton $\{0\}$ and $(k-1)/2$ orbits of type $\{\pm a\}$, $a \neq 0$ in $\mathbb{Z}_k$, with cardinality 2. The homothecy $\vartheta_2 : \mathbb{Z}_k \to \mathbb{Z}_k$ induces a permutation $\overline{\vartheta}_2$ of $\overline{X}_k$. To compute the order of $\overline{\vartheta}_2$ in $\mathfrak{S}(\overline{X}_k)$, we need to determine the $\ell \in \mathbb{Z}$ with $\{\pm 2^{\ell} a\} = \{\pm a\}$ for all $a \in \mathbb{Z}_k$. This is equivalent with $\{\pm 2^{\ell}\} = \{\pm 1\}$ or with $2^{\ell} \in \{\pm 1\}$. If $-1 \notin \mathrm{H}(2)$, then necessarily $2^{\ell} = 1$ and $\ell \in \mathbb{Z}\operatorname{ord}_k 2$. If $-1 \in \mathrm{H}(2) \subseteq \mathbb{Z}_k^{\times}$, then the order $\operatorname{ord}_k 2$ of the cyclic group $\mathrm{H}(2)$ is even and $2^{\frac{1}{2}\operatorname{ord}_k 2} = -1$ and $\ell \in \mathbb{Z}\frac{1}{2}\operatorname{ord}_k 2$. Altogether,

$$\operatorname{ord}\overline{\vartheta}_2 = \begin{cases} \operatorname{ord}_k 2, & \text{if } -1 \notin \mathrm{H}(2) \subseteq \mathbb{Z}_k^{\times}, \\ \frac{1}{2}\operatorname{ord}_k 2, & \text{if } -1 \in \mathrm{H}(2) \subseteq \mathbb{Z}_k^{\times}. \end{cases}$$

Now, to compute the order $\operatorname{ord}\sigma_{\mathrm{MB}}$ in the even case $n = 2m$, we interpret the elements $1, 2, \ldots, 2m$ as representatives of the orbits of $\overline{X}'_{4m+1} := \overline{X}_{4m+1} \backslash \{\{0\}\}$. Then σ_{MB} is the permutation $\overline{\vartheta}_2 | \overline{X}'_{4m+1}$. For $\operatorname{ord}\sigma_{\mathrm{MA}}$, we omit the fixed point 1 and translate the elements $2, \ldots, 2m$ by -1 and get the permutation

$$\sigma'_{\mathrm{MA}} = \begin{pmatrix} 1 & 2 & \ldots & m-1 & m & m+1 & \ldots & 2m-2 & 2m-1 \\ 2 & 4 & \ldots & 2m-2 & 2m-1 & 2m-3 & \ldots & 3 & 1 \end{pmatrix},$$

which we interpret as $\overline{\vartheta}_2 | \overline{X}'_{4m-1}$. Analogously, in the odd case $n = 2m+1$, we interpret σ_{MB} and σ_{MA} as $\overline{\vartheta}_2 | \overline{X}'_{4m+3}$ and $\overline{\vartheta}_2 | \overline{X}'_{4m+1}$, respectively. Using the above general formula for $\operatorname{ord}\overline{\vartheta}_2$, the proof of 3.7 is complete.

We mention the following corollary which will be used in Example 3.9.

3.8 Corollary (1) *Let* $n = 2m \in \mathbb{N}^*$ *be even. For the Monge permutation* $\sigma_{\mathrm{MB}} \in \mathfrak{S}_n$*, the following statements are equivalent* :

(i) σ_{MB} *is a cycle of order* $n = 2m$.

(ii) $\operatorname{ord}\sigma_{\mathrm{MB}} = n = 2m$.

(iii) $p := 2n+1 = 4m+1$ *is prime and* $\operatorname{ord}_p 2 = p-1$.

(2) *Let* $n = 2m+1 \in \mathbb{N}^*$ *be odd. For the Monge permutation* $\sigma_{\mathrm{MB}} \in \mathfrak{S}_n$*, the following statements are equivalent* :

(i) σ_{MB} *is a cycle of order* $n = 2m+1$.

(ii) $\operatorname{ord} \sigma_{\mathrm{MB}} = n = 2m+1$.

(iii) $p := 2n+1 = 4m+3$ *is prime and either* $\operatorname{ord}_p 2 = (p-1)/2$ *or* $\operatorname{ord}_p 2 = p-1$.

Note that the permutation σ_{MA} of a Monge out-shuffling is never a cycle of order $n = \#C$ if $n > 1$, since 1 is a fixed point of σ_{MA}.

To prove the implication "(ii) $\Rightarrow$ (iii)" in part (1) of Corollary 3.8 assume that $\operatorname{ord} \sigma_{\mathrm{MB}} = n = 2m$ and $-1 \notin \mathrm{H}(2) \subseteq \mathbb{Z}_{4m+1}$. Then, by Theorem 3.7(1), $\operatorname{ord}_{4m+1} 2 = 2m$ and $\#\big(\mathrm{H}(2) \cdot \{\pm 1\}\big) = 4m$, hence $p := 4m+1$ is prime and $\mathbb{Z}_p^\times \cong \mathrm{H}(2) \times \{\pm 1\}$. This contradicts the cyclicity of the group $\mathbb{Z}_p^\times$.[33] It follows $-1 \in \mathrm{H}(2)$ and, again by Theorem 3.7 (1), $\operatorname{ord}_{4m+1} 2 = 4m$, hence $p = 4m+1$ is prime and $\operatorname{ord}_p 2 = p-1$. The implication "(iii) $\Rightarrow$ (i)" follows from the interpretation of σ_{MB} in the proof of Theorem 3.7.

To prove the implication "(ii) $\Rightarrow$ (iii)" in part (2) of Corollary 3.8 assume that $\operatorname{ord} \sigma_{\mathrm{MB}} = n = 2m+1$ and $-1 \notin \mathrm{H}(2) \subseteq \mathbb{Z}_{4m+3}$. Then, by Theorem 3.7(2), $\operatorname{ord}_{4m+3} 2 = 2m+1$ and $\#\big(\mathrm{H}(2) \cdot \{\pm 1\}\big) = 4m+2$, hence $p := 4m+3$ is prime and $\operatorname{ord}_p 2 = (p-1)/2$. If $-1 \in \mathrm{H}(2)$, then, cf. Theorem 3.7(2), $\operatorname{ord}_{4m+3} 2 = 4m+2$ and $p = 4m+3$ is prime with $\operatorname{ord}_p 2 = p-1$. The implication "(iii) $\Rightarrow$ (i)" follows again from the proof of Theorem 3.7.

We mention that, if the statements of Corollary 3.8 (1) are true, then, by Proposition 2.11.5, $p = 4m+1 \equiv 3$ or $5 \bmod 8$ and hence even $p = 4m+1 \equiv 5 \bmod 8$. If the statements of Corollary 3.8 (2) are true, then $p = 4m+3 \equiv 7 \bmod 8$ if $\operatorname{ord}_p 2 = (p-1)/2$ and $p \equiv 3 \bmod 8$ if $\operatorname{ord}_p 2 = p-1$. In particular, p is never $\equiv 1 \bmod 8$. If $(n = 2m+1, p = 2n+1 = 4m+3)$ is a Sophie Germain pair of odd primes n, p, cf. Example 2.8, then the statements of Corollary 3.8 (2) are true.

3.9 Example (Monge shufflings in poetry) A sestina is a poem of six six-line stanzas in which each stanza repeats the end words of the lines of the previous stanza according to the Monge in-shuffling

$$\sigma_{\mathrm{MB}} := \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 6 & 5 & 3 & 1 \end{pmatrix} = \langle 1,2,4,5,3,6 \rangle \in \mathfrak{S}_6$$

(supplemented with a three-line envoy which we ignore here).[34] If one wants to form a poem in an analogous manner as a sestina, but with n stanzas each consisting of n lines, $n \in \mathbb{N}^*$, then the corresponding Monge permutation $\sigma_{\mathrm{MB}} \in \mathfrak{S}_n$ should be an n-cycle so that each end word of the lines of the first stanza will run in the following stanzas through all the lines and, moreover, the shuffling of the end words of the lines of the last stanza will yield the end words of the lines of the first stanza in their correct order. The problem to find those n was already posed and (partially) solved by the French writer Raymond Queneau (1903 - 1976). By Corollary 3.8, $n \in \mathbb{N}^*$ *is an appropriate number if and only if it is even and fulfills condition* (iii) *in* Corollary 3.8(1) *or it is odd and fulfills condition* (iii) *in* Corollary 3.8 (2). The even numbers $n \leq 30$ of this kind are 2, 6, 14, 18, 26, 30, and the odd numbers $n \leq 30$ of this kind are 1, 3, 5, 9, 11, 23, 29.

[33] Or, more elementary: The equation $x^2 = 1$ has in $\mathrm{H}(2) \times \{\pm 1\}$ four solutions (since $\#\mathrm{H}(2)$ is even) but in $\mathbb{Z}_p^\times$ only two.

[34] The sestina was (probably) invented by the Provençal troubadour Arnaut Daniel in the second half of the 12th century. Later on it played a significant role in Italian poetry. – Sometimes the simple cycle $\gamma_6 = \langle 1,2,3,4,5,6 \rangle$ as in Example 1.1.1 is used to shuffle the end words of the lines of a stanza. – The envoy repeats the end words of the first stanza, three in the middle and three at the end of its lines. An example of a sestina can be found in Math. Int. **49** (1), p. 7 (2007) . For an original sestina by Arnaut Daniel cf. the book "Prosody in England and Elsewhere: A Comparative Approach" by L. Malcovati.

§ 4 Shuffling of Multisets

For many card games the set C of cards is a multiset. It contains cards which are indistinguishable. For example, Canasta is played with two decks of 52 cards each and 4 jokers. Therefore, the whole set of 108 cards contains 52 pairs of two indistinguishable cards each and one set of 4 indistinguishable jokers.

4.1 Multisets For our purposes, the best method to model this situation is to define a multiset $(C;\mathcal{R})$ as a set C together with an equivalence relation $\mathcal{R}$ on C given by a subset $\mathcal{R} \subseteq \mathfrak{P}(C)$ which is a partition of C into non-empty subsets. For the automorphism group $\mathfrak{S}(C;\mathcal{R})$ of the multiset $(C;\mathcal{R})$ one chooses the subgroup $\prod_{R\in\mathcal{R}} \mathfrak{S}(R) \subseteq \mathfrak{S}(C)$ of those permutations $\rho \in \mathfrak{S}(C)$ which map each set $R \in \mathcal{R}$ into itself, i. e.

$$\mathfrak{S}(C;\mathcal{R}) := \{\rho \in \mathfrak{S}(C) \mid \rho(R) = R \text{ for every } R \in \mathcal{R}\}.$$

Two stacks $\mathbf{c},\mathbf{d} : [1,n] \to C$ in $\mathfrak{P}_C$ ($n := \#C < \infty$) are indistinguishable if $\mathbf{d} = \rho\,\mathbf{c} (= \rho \circ \mathbf{c})$ for some $\rho \in \mathfrak{S}(C;\mathcal{R})$. Hence, the set of multipacks of the multiset $(C;\mathcal{R})$ is the set

$$\mathfrak{P}_{(C;\mathcal{R})} := \mathfrak{P}_C \backslash \mathfrak{S}(C;\mathcal{R})$$

of orbits $[\mathbf{c}] = [\mathbf{c}]_{\mathcal{R}} := \mathfrak{S}(C;\mathcal{R})\mathbf{c}$ of the canonical (left) operation of $\mathfrak{S}(C;\mathcal{R}) \subseteq \mathfrak{S}(C) = \mathfrak{S}(C) \times \{\mathrm{id}_n\}$ on $\mathfrak{P}_C$ described already at the end of Example 1.6. The cardinality of $\mathfrak{P}_{(C;\mathcal{R})}$ is $[\mathfrak{S}(C) : \mathfrak{S}(C;\mathcal{R})] = n!/\prod_{R\in\mathcal{R}}(\#R)!$.[35] Since the operations of $\mathfrak{S}(C)$ and $\mathfrak{S}_n$ on $\mathfrak{P}_C$ commute, *the shuffling induces an operation*

$$\mathfrak{S}_n \times \mathfrak{P}_{(C;\mathcal{R})} \to \mathfrak{P}_{(C;\mathcal{R})},\ (\sigma,[\mathbf{c}]) \mapsto \sigma * [\mathbf{c}] = [\sigma * \mathbf{c}] = [\mathbf{c}\sigma^{-1}],$$

of $\mathfrak{S}_n$ *on* $\mathfrak{P}_{(C;\mathcal{R})}$ *which is again transitive*, but *not* simply transitive (if $\mathcal{R}$ is not the identity relation). On the contrary, the stabilizer of the stack $[\mathbf{c}]_{\mathcal{R}}$ is the group

$$\mathfrak{S}_{n,[\mathbf{c}]_{\mathcal{R}}} = \{\sigma \in \mathfrak{S}_n \mid [\mathbf{c}\sigma^{-1}]_{\mathcal{R}} = [\mathbf{c}]_{\mathcal{R}}\} = \{\mathbf{c}^{-1}\rho\mathbf{c} \mid \rho \in \mathfrak{S}(C;\mathcal{R})\},$$

i. e. the conjugate $\mathbf{c}^{-1}\mathfrak{S}(C;\mathcal{R})\mathbf{c}$ of $\mathfrak{S}(C;\mathcal{R})$ in $\mathfrak{S}_n$ with respect to the bijective map $\mathbf{c}^{-1} : C \to [1,n]$. Note that $\mathbf{c}^{-1}\mathfrak{S}(C;\mathcal{R})\,\mathbf{c} = \mathfrak{S}([1,n];\mathbf{c}^{-1}(\mathcal{R})) = \prod_{r\in\mathcal{R}} \mathfrak{S}(\mathbf{c}^{-1}(R)) \subseteq \mathfrak{S}_n$. For a fixed shuffling permutation $\sigma \in \mathfrak{S}_n$ the stacks $[\mathbf{c}]_i := \sigma^i * [\mathbf{c}]_{\mathcal{R}}$ and $[\mathbf{c}]_j := \sigma^j * [\mathbf{c}]_{\mathcal{R}} = \sigma^{j-i} * [\mathbf{c}]_i$ coincide for $i,j \in \mathbb{Z}$ if and only if $\sigma^{j-i} \in \mathrm{H}(\sigma) \cap \mathfrak{S}_{n,[\mathbf{c}]_{\mathcal{R}}}$, i. e. if and only if

$$i \equiv j \bmod \frac{\operatorname{ord}\sigma}{\#\big(\mathrm{H}(\sigma) \cap \mathbf{c}^{-1}\mathfrak{S}(C;\mathcal{R})\,\mathbf{c}\big)}.$$

This proves the following theorem:

4.2 Theorem *The sequence* $[\mathbf{c}]_0 = [\mathbf{c}]_{\mathcal{R}}, [\mathbf{c}]_1 = \sigma * [\mathbf{c}]_{\mathcal{R}}, [\mathbf{c}]_2 = \sigma^2 * [\mathbf{c}]_{\mathcal{R}} = \sigma * [\mathbf{c}]_1, \ldots,$

[35] A similar situation occurs for a pair $(C;\mathcal{R})$ where C is a set of n not necessarily distinct characters and, where the equivalence classes $R \in \mathcal{R}$ are the sets of identical characters respectively. $\mathfrak{P}_{(C;\mathcal{R})}$ is then the set of the different words of length n which can be formed with the characters contained in C.

$[\mathbf{c}]_{i+1} = \sigma^{i+1} * [\mathbf{c}]_{\mathcal{R}} = \sigma * [\mathbf{c}]_i, \ldots$ *is (purely) periodic of period length*

$$\frac{\operatorname{ord}\sigma}{\#\big(\mathrm{H}(\sigma)\cap \mathbf{c}^{-1}\mathfrak{S}(C;\mathcal{R})\,\mathbf{c}\big)}\,.$$

Note that the period length depends significantly on the stack $[\mathbf{c}]_{\mathcal{R}}$ which the shuffling starts with. To determine the possible orders $\#\big(\mathrm{H}(\sigma)\cap \mathbf{c}^{-1}\mathfrak{S}(C;\mathcal{R})\,\mathbf{c}\big)$, $\mathbf{c}\in\mathfrak{P}_C$, one usually needs the cycle decomposition of σ, cf. Example 1.7.

To describe this more precisely, let us first introduce the following terminology for partitions: If $\mathcal{Y}$ and $\mathcal{Z}$ are partitions of the same set X, then $\mathcal{Y}$ is called `finer` than $\mathcal{Z}$ if every $Y\in\mathcal{Y}$ is contained in some $Z\in\mathcal{Z}$ or, equivalently, if every $Z\in\mathcal{Z}$ is the union of elements of $\mathcal{Y}$. Similarly, if $n=\sum_{i\in I} y_i = \sum_{j\in J} z_j$ are partitions of $n\in\mathbb{N}$ with positive integers y_i, z_j, then $\sum_{i\in I} y_i$ is called `finer` than $\sum_{j\in J} z_j$ if there is a partition $I=\biguplus_{j\in J} I_j$ of the index set I with $z_j=\sum_{i\in I_j} y_i$ for all $j\in J$.

Generally, a permutation $\tau\in\mathfrak{S}(X)$ of a finite set X belongs to $\mathfrak{S}(X;\mathcal{Z})$ where $\mathcal{Z}$ is a partition of X if and only if the partition of X defined by the orbits of τ is finer than $\mathcal{Z}$. We obtain: A permutation $\tau\in\mathfrak{S}_n$ belongs to $\mathbf{c}^{-1}\mathfrak{S}(C;\mathcal{R})\,\mathbf{c}$ for some $\mathbf{c}\in\mathfrak{P}_C$ if and only if the partition $1^{\nu_1(\tau)}2^{\nu_2(\tau)}\ldots n^{\nu_n(\tau)}$ of $n=\#C$ defined by the type of τ is finer than the partition $1^{\nu_1(\mathcal{R})}2^{\nu_2(\mathcal{R})}\ldots n^{\nu_n(\mathcal{R})}$ of n defined by $\mathcal{R}$ (with $\nu_i(\mathcal{R}) := \#\{R\in\mathcal{R}\mid \#R=i\}$). Now, to decide whether $\sigma^k\in\mathbf{c}^{-1}\mathfrak{S}(C;\mathcal{R})\,\mathbf{c}$ for some $k\in\mathbb{N}^*$ and some $\mathbf{c}\in\mathfrak{P}_C$ one uses the formula

$$\nu_i(\sigma^k)\cdot i = \sum_{j,\, i\cdot \mathrm{GCD}(j,k)=j} \nu_j(\sigma)\cdot j\,, \qquad i,k\in\mathbb{N}^*,$$

which follows from the fact that the k-th power of a cycle of length $m\in\mathbb{N}^*$ is a product of $\mathrm{GCD}(m,k)$ cycles of length $m/\mathrm{GCD}(m,k)$.

4.2.1 Example For the Old Maid game (cf. Example 3.3.4) the partition $\mathcal{R}$ of C contains 16 pairs and a singleton. Hence, in this case, $\mathfrak{S}(C;\mathcal{R})$ is an elementary 2-group of order 2^{16} and every element of order 2 in $\mathfrak{S}_{33}$ belongs to some $\mathbf{c}^{-1}\mathfrak{S}(C;\mathcal{R})\,\mathbf{c}$. For $\sigma\in\mathfrak{S}_{33}$ and $\mathbf{c}\in\mathfrak{P}_C$ the group $\mathrm{H}(\sigma)\cap\mathbf{c}^{-1}\mathfrak{S}(C;\mathcal{R})\,\mathbf{c}$ is of order 1 or 2. It follows: If $\operatorname{ord}\sigma$ is even *and* $[\sigma^{\operatorname{ord}\sigma/2}*\mathbf{c}]=[\mathbf{c}]$, then the period length of the sequence $[\mathbf{c}]_i=\sigma^i*[\mathbf{c}]$, $i\in\mathbb{N}$, is $\frac{1}{2}\operatorname{ord}\sigma$. In all the other cases the period length is $\operatorname{ord}\sigma$. – For the Monge shufflings one has, by Theorem 3.7, $\operatorname{ord}\sigma_{\mathrm{MA}}=\frac{1}{2}\operatorname{ord}_{65}2=6$ (since $\operatorname{ord}_{65}2=\mathrm{LCM}(\operatorname{ord}_5 2,\operatorname{ord}_{13}2)=\mathrm{LCM}(4,12)=12$ and $-1\equiv 2^6 \mod 65$) and $\operatorname{ord}\sigma_{\mathrm{MB}}=\frac{1}{2}\operatorname{ord}_{67}2=33$ (since $2^6\equiv -3 \mod 67$, $2^{33}\equiv -1 \mod 67$ by Proposition 2.11.5, $2^{22}=(2^6)^3\cdot 2^4\equiv(-3)^3\cdot 2^4\equiv 40\cdot 2^4=2^6\cdot 10\equiv -30 \mod 67$ and hence $\operatorname{ord}_{67}2=66$). Hence, for the Monge out-shuffling $\sigma_{\mathrm{MA}}\in\mathfrak{S}_{33}$ the period length may be 6 or 3, and for the Monge in-shuffling $\sigma_{\mathrm{MB}}\in\mathfrak{S}_{33}$ it is always 33 (independent of the original stack $[\mathbf{c}]$). By Corollary 3.8(2) σ_{MB} is a cycle of length 33.

4.2.2 Example For Canasta as described at the beginning of Section 4 the partition $\mathcal{R}$ of C contains 52 pairs and 1 quartet (four jokers). Let us consider the Faro shufflings $\sigma_{1\mathrm{A}},\sigma_{1\mathrm{B}}\in\mathfrak{S}_{108}$. By Theorem 3.4(1), one has $\operatorname{ord}\sigma_{1\mathrm{A}}=\operatorname{ord}_{107}2=106$ and $\operatorname{ord}\sigma_{1\mathrm{B}}=\operatorname{ord}_{109}2=36$ (cf. Example 3.4.2). It follows that $\#\big(\mathrm{H}(\sigma_{1\mathrm{A}})\cap\mathfrak{S}(C;\mathcal{R})\big)\le 2$. Further, we have

$$\sigma_{1\mathrm{A}}^{53}=\begin{pmatrix}1 & 2 & 3 & \ldots & 106 & 107 & 108\\ 1 & 107 & 106 & \ldots & 3 & 2 & 108\end{pmatrix}.$$

The period length of the Faro out-shuffling $\sigma_{1A} \in \mathfrak{S}_{108}$ is 53 if in the original stack $\mathbf{c}=(c_1,\dots,c_{108})$ the pairs $\{c_2,c_{107}\},\dots,\{c_{54},c_{55}\}$ contain indistinguishable cards each, otherwise the period length is 106. – The permutation $\sigma_{1B} \in \mathfrak{S}_{108}$ can be interpreted as multiplication by 2 in $\mathbb{Z}_{109}^{\times}$ and hence its cycle decomposition contains 3 cycles of length 36. Again, it follows $\#\big(\mathrm{H}(\sigma_{1B}) \cap \mathfrak{S}(C;\mathcal{R})\big) \le 2$. Since

$$\sigma_{1B}^{18} = \begin{pmatrix} 1 & 2 & 3 & \dots & 106 & 107 & 108 \\ 108 & 107 & 106 & \dots & 3 & 2 & 1 \end{pmatrix}$$

is multiplication by $2^{18} = -1$ in $\mathbb{Z}_{109}^{\times}$, *the period length of the Faro in-shuffling σ_{1B} is 18 if in the original stack $\mathbf{c} = (c_1,\dots,c_{108})$ the pairs $\{c_1,c_{108}\},\dots,\{c_{54},c_{55}\}$ contain indistinguishable cards each, and 36 otherwise.*

4.2.3 Example (Four-games) In a four-game (in German Quartett, in English often called Happy Families) the partition $\mathcal{R}$ of the multiset $(C;\mathcal{R})$ contains k sets with four elements each[36], hence $\#C = n = 4k$, and the group $\mathfrak{S}(C;\mathcal{R})$ is isomorphic to $\mathfrak{S}_4^k$. Let us consider the Faro shufflings with $\sigma_{1A}, \sigma_{1B} \in \mathfrak{S}_{32}$ in case $n = 32$, i. e. $k = 8$. By Theorem 3.4,

$$\operatorname{ord} \sigma_{1A} = \operatorname{ord}_{31} 2 = 5\,, \quad \operatorname{ord} \sigma_{1B} = \operatorname{ord}_{33} 2 = 10\,.$$

Therefore $\#\big(\mathrm{H}(\sigma_{1A}) \cap \mathbf{c}^{-1}\mathfrak{S}(C;\mathcal{R})\mathbf{c}\big) = 1$ and $\#\big(\mathrm{H}(\sigma_{1B}) \cap \mathbf{c}^{-1}\mathfrak{S}(C;\mathcal{R})\mathbf{c}\big) \le 2$ for all $\mathbf{c} \in \mathfrak{P}_C$. Hence, by Theorem 4.2, the period length for the Faro out-shuffling is always 5, and for the Faro in-shuffling it is 10 or 5 depending on the original pack $\mathbf{c}$. The reader may also consider the case $n = 40$, i. e. $k = 10$.

4.3 Remark So far we have not taken into account the role of the dealer of the cards. Each player is dealt a certain number of cards and the rest (if any) forms a stock which may be an ordered pile or an unordered set. For example, in Canasta the four players are dealt 11 cards each and the rest of 64 cards is placed as an ordered stack pile. In the European card game Skat, each of the three players is dealt 10 cards and 2 additional cards form an unordered stock (called the "Skat"). Usually the dealer distributes the cards according to a fixed scheme, i.e. each of the players gets a set $\mathbf{c}(T)$ of the stack $\mathbf{c} : [1,n] \to C$ where the subset $T \subseteq [1,n]$ is fixed. The game does not change if the stack $\mathbf{c}$ is replaced by $\sigma * \mathbf{c} = \mathbf{c}\sigma^{-1}$ for any $\sigma \in \mathfrak{S}(T) \subseteq \mathfrak{S}_n$. This means, one has to consider besides the partition $\mathcal{R}$ of the set C of cards a partition $\mathcal{T}$ of the set $[1,n]$ of the positions of the cards in a stack. For example, for Canasta the partition $\mathcal{T}$ (of $[1,108]$) contains 4 sets of 11 elements each and 64 singletons, for Skat $\mathcal{T}$ contains 3 sets of 10 elements each and 1 pair. In Subsection 4.1 we discussed the case that $\mathcal{T}$ is the identity relation on $[1,n]$, i. e. that $\mathcal{T}$ contains only singletons.

Two stacks $\mathbf{c},\mathbf{d} \in \mathfrak{P}_C$ have to be identified if there are a $\rho \in \mathfrak{S}(C;\mathcal{R}) = \prod_{R\in\mathcal{R}} \mathfrak{S}(R)$ and a $\tau \in \mathfrak{S}([1,n];\mathcal{T}) = \prod_{T\in\mathcal{T}} \mathfrak{S}(T)$ with $\mathbf{d} = (\rho,\tau) * \mathbf{c} = \rho\,\mathbf{c}\,\tau^{-1}$, cf. the end of Example 1.6. In other words, the multistacks one has to consider now are the "double cosets"

$$[\![\mathbf{c}]\!] = [\![\mathbf{c}]\!]_{\mathcal{R},\mathcal{T}} = (\mathfrak{S}(C;\mathcal{R}) \times \mathfrak{S}([1,n];\mathcal{T})) * \mathbf{c} = \mathfrak{S}(C;\mathcal{R})\,\mathbf{c}\,\mathfrak{S}([1,n];\mathcal{T})\,.$$

These are the orbits of the canonical (left) operation of $\mathfrak{S}(C;\mathcal{R}) \times \mathfrak{S}([1,n];\mathcal{T}) \subseteq \mathfrak{S}(C) \times \mathfrak{S}_n$ on $\mathfrak{P}_C$. We denote the set of these orbits by $\mathfrak{P}_{(C;\mathcal{R},\mathcal{T})} = \mathfrak{P}_{\mathfrak{S}(C;\mathcal{R})} \backslash \mathfrak{S}([1,n];\mathcal{T})$. The stabilizer of $\mathbf{c} \in \mathfrak{P}_C$ is the subgroup

$$\begin{aligned} \{(\rho,\tau) \in \mathfrak{S}(C;\mathcal{R}) \times \mathfrak{S}([1,n];\mathcal{T}) \mid \rho = \mathbf{c}\tau\mathbf{c}^{-1}\} &\cong \mathfrak{S}(C;\mathcal{R}) \cap \mathfrak{S}(C;\mathbf{c}(\mathcal{T})) \\ &= \mathfrak{S}(C;\mathcal{R} \sqcap \mathbf{c}(\mathcal{T}))\,, \end{aligned}$$

where, for two partitions $\mathcal{Y},\mathcal{Z}$ of a set X, the partition $\mathcal{Y} \sqcap \mathcal{Z}$ contains the non-empty intersections

[36] To be correct, the cards of a four-set $R \in \mathcal{R}$ are usually distinguishable. But, in general, this plays no essential role in the course of the game.

$Y \cap Z$ with $Y \in \mathcal{Y}, Z \in \mathcal{Z}$. In particular (cf. Theorem 1.3),

$$\#[\![\mathbf{c}]\!]_{\mathcal{R},\mathcal{T}} = \frac{\#\mathfrak{S}(C;\mathcal{R}) \cdot \#\mathfrak{S}([1,n];\mathcal{T})}{\#\mathfrak{S}(C;\mathcal{R} \sqcap \mathbf{c}(\mathcal{T}))}.$$

In general, this number depends not only on $\mathcal{R}$ and $\mathcal{T}$, but also on $\mathbf{c}$.

Now choose a fixed shuffling permutation $\sigma \in \mathfrak{S}_n$, $n = \#C$. The sequence of multistacks obtained from $[\![\mathbf{c}]\!]_{\mathcal{R},\mathcal{T}}$ by iterated shuffling with σ is

$$[\![\mathbf{c}]\!]_i := [\![\sigma^i * \mathbf{c}]\!]_{\mathcal{R},\mathcal{T}} = \mathfrak{S}(C;\mathcal{R})\,\mathbf{c}\,\sigma^{-i}\mathfrak{S}([1,n];\mathcal{T}),\ i \in \mathbb{N}.$$

The difficulty in studying this sequence arises from the fact that, in general, the operation of $\mathfrak{S}_n$ on $\mathfrak{P}_C$ does *not* induce an operation of $\mathfrak{S}_n$ on $\mathfrak{P}_{(C;\mathcal{R},\mathcal{T})}$, since, for $\tau \in \mathfrak{S}_n$ and $\mathbf{d}, \mathbf{d}' \in \mathfrak{P}_C$, the equality $[\![\mathbf{d}]\!]_{\mathcal{R},\mathcal{T}} = [\![\mathbf{d}']\!]_{\mathcal{R},\mathcal{T}}$ does *not* necessarily imply $[\![\tau * \mathbf{d}]\!]_{\mathcal{R},\mathcal{T}} = [\![\tau * \mathbf{d}']\!]_{\mathcal{R},\mathcal{T}}$.

To treat this problem in a more general setting, let X be a G-space $H \subseteq G$ a subgroup and consider the induced operation of H on X. Then, an element $g \in G$ operates naturally on the set $X\backslash H$ of H-orbits if, for every $x \in X$, the set gHx coincides with the H-orbit Hgx of gx. The subgroup of the elements with this property contains the normalizer $\mathrm{N}_G(H) = \{g \in G \mid gH = Hg\}$ of H in G. In particular, *G operates canonically on $X\backslash H$ if H is a normal subgroup of G.* In this case the projection $X \to X\backslash H$ is a G-morphism, and the induced mapping $X\backslash G \to (X\backslash H)\backslash G$ is bijective.

In our situation, X is the set $\mathfrak{P}_C$ with the operation of $G := \mathfrak{S}(C) \times \mathfrak{S}_n$ and $H \subseteq G$ is the subgroup $H := \mathfrak{S}(C;\mathcal{R}) \times \mathfrak{S}([1,n];\mathcal{T})$. The normalizer $\mathrm{N}_G(H)$ is the group

$$\mathfrak{S}(C)_{\mathcal{R}} \times \mathfrak{S}_{n,\mathcal{T}} = \{(\rho,\tau) \in \mathfrak{S}(C) \times \mathfrak{S}_n \mid \rho(\mathcal{R}) = \mathcal{R}, \tau(\mathcal{T}) = \mathcal{T}\}$$

(because of $(\rho,\tau)H(\rho^{-1},\tau^{-1}) = \mathfrak{S}(C;\rho(\mathcal{R})) \times \mathfrak{S}([1,n];\tau(\mathcal{T}))$). The sequence $[\![\mathbf{c}]\!]_i = [\![\mathbf{c}\,\sigma^{-i}]\!]_{\mathcal{R},\mathcal{T}}$, $i \in \mathbb{N}$, is purely periodic, and its period length divides ord σ (cf. Footnote 25). Since $[\![\mathbf{c}]\!]_{i+j} = [\![\mathbf{c}]\!]_i$ is equivalent with $\sigma^j \in \mathfrak{S}([1,n];\sigma^i(\mathcal{T})) \cdot \mathfrak{S}([1,n];\mathbf{c}^{-1}(\mathcal{R}))$, *the period length of the purely periodic sequence* $([\![\mathbf{c}]\!]_i)_{i\in\mathbb{N}}$ *is*

$$\ell(\mathbf{c}) := \ell_{\mathcal{R},\mathcal{T}}(\mathbf{c}) := \frac{\text{ord } \sigma}{\#\bigcap_{i=0}^{\text{ord }\sigma-1} \mathrm{H}(\sigma) \cap (\mathfrak{S}([1,n];\sigma^i(\mathcal{T})) \cdot \mathfrak{S}([1,n];\mathbf{c}^{-1}(\mathcal{R})))}.$$

Note that

$$\bigcap_{i=0}^{\text{ord }\sigma-1} \mathrm{H}(\sigma) \cap \big(\mathfrak{S}([1,n];\sigma^i(\mathcal{T})) \cdot \mathfrak{S}([1,n];\mathbf{c}^{-1}(\mathcal{R}))\big)$$

contains the group $(\mathrm{H}(\sigma) \cap \mathfrak{S}([1,n];\mathcal{T})) \cdot \big(\mathrm{H}(\sigma) \cap \mathfrak{S}([1,n];\mathbf{c}^{-1}(\mathcal{R}))\big)$ of order

$$\mathrm{LCM}\left(\#(\mathrm{H}(\sigma) \cap \mathfrak{S}([1,n];\mathcal{T})), \#(\mathrm{H}(\sigma) \cap \mathfrak{S}([1,n];\mathbf{c}^{-1}(\mathcal{R})))\right).$$

We emphasize that the elements $[\![\mathbf{c}]\!] = [\![\mathbf{c}]\!]_0, \ldots, [\![\mathbf{c}]\!]_{\ell(\mathbf{c})-1}$ of the period of the sequence $([\![\mathbf{c}]\!]_i)_{i\in\mathbb{N}}$ are not necessarily pairwise distinct.

References

[1] Conway, J. H.; Guy, R. K.: *The book of numbers*, Springer, New York 1996.

[2] Gardner, M.: *Mathematical Carnival*, Alfred A. Knopf, New York 1975.

[3] Gauss, C. F.: *Disquisitiones arithmeticae*, Leipzig 1801. English translation by Clarke, A. A., Waterhouse, W. C., Greither, C. and Grootendorst, A. W., Springer, New York 1986.

[4] Knuth, D. E.: *The Art of Computer Programming*, Vol. 2. Addison-Wesley, Reading Mass. ^{3}1998.

[5] Scheja, G.; Storch, U.: *Lehrbuch der Algebra*, Teil 1, 2. B. G. Teubner, Stuttgart ^{2}1994, 1988.

[6] Storch, U.: *Pythagoras and Diophantus*, Resonance (Journal of Science Education), **14** (2009), No. 7, 691-703.

[7] Storch, U.; Wiebe, H.: *Lehrbuch der Mathematik* Band. 1, 2, 3, 4. Spektrum Akademischer Verlag, Heidelberg ^{3}2009, ^{2}2010, 2010, 2011.

Received 8 November 2010; Revised 10 January 2011

D i l i p P. P a t i l was born in Anjale village of Jalgaon district (Maharashtra) on May 02, 1956. His primary and secondary education was completed in Anjale. He then completed B. Sc. in Mathematics from Abasaheb Garware College, Pune in 1976 and M. Sc. in Mathematics from University of Pune in 1978. From 1978 till 1992 he studied Mathematics at the School of Mathematics, Tata Institute of Fundamental Research, Mumbai and received Ph. D. through University of Bombay in 1989 for his thesis related to defining equations of algebraic curves. He joined the Department of Mathematics at IISc in 1992 and is currently a Professor there as well as in the Department of Computer Science and Automation, IISc, Bangalore.

He has been a recipient of the C. L. Chandna Mathematics Award for the year 2001 (presented by Canadian World Education Foundation Saraswati Vishvas, Canada) in recognition of distinguished contributions to Mathematics Research and Teaching. He has been a Visiting Professor at Ruhr- Universität Bochum, Universität Leipzig and several universities in Europe and Canada. His research interests are mainly in Commutative Algebra and Algebraic Geometry.

U w e S t o r c h studied Mathematics, Physics and Mathematical Logic at the Universität Münster and Heidelberg from 1960 till 1966 and received Ph. D. from Universität Münster in 1966. In 1972 Habilitation at the Ruhr-Universität Bochum. From 1974 till 1981 and from 1981 till 2005 he was a Full Professor at the Universität Osnabrück and at the Ruhr-Universität Bochum, respectively, holding chairs on Algebra and Geometry. Currently he is Professor Emeritus at the Ruhr-Universität Bochum. He has been a Visiting Professor at Tata Institute of Fundamental Research, Mumbai, Indian Institute of Science, Bangalore and several universities in Europe and USA. His research interests are mainly in Algebra, particularly the algebraic aspects of Complex Analytic Geometry, Commutative Algebra and Algebraic Geometry.