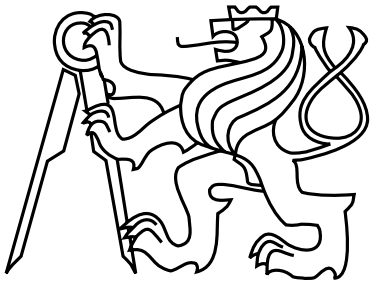
Czech Technical University in Prague
Faculty of Electrical Engineering
Department of Cybernetics



Master's Thesis

Using Monodromy to Simplify Polynomial Systems

Bc. Viktor Korotynskiy

Supervisor: doc. Ing. Tomáš Pajdla, Ph.D.

Study Program: Robotics, Master

Field of Study: Robotics

May 22, 2020

MASTER‘S THESIS ASSIGNMENT

I. Personal and study details

Student's name: **Korotynskiy Viktor** Personal ID number: **453214**

Faculty / Institute: **Faculty of Electrical Engineering**

Department / Institute: **Department of Cybernetics**

Study program: **Cybernetics and Robotics**

Branch of study: **Robotics**

II. Master’s thesis details

Master’s thesis title in English:

Using Monodromy to Simplify Polynomial Systems

Master’s thesis title in Czech:

Použití monodromie ke zjednodušení polynomiálních soustav

Guidelines:

1) Review simplification of polynomial systems, Galois theory and relevant parts of algebraic topology related to monodromy of polynomial systems.
2) Explain the connection between the monodromy group of a polynomial system with a finite number of solutions, its Galois group and symmetries of this polynomial system.
3) Find an example of a relevant polynomial system that can be simplified by using its monodromy group.

Bibliography / sources:

[1] A. Hatcher. Algebraic Topology. Cambridge University Press, 2002.
[2] J. Harris. Galois groups of enumerative problems. Duke Math. Journal, 1979.
[3] C. Améndola, J. Rodriguez. Solving parametrized polynomial systems with decomposable projections. arXiv:1612.08807 (2016).
[4] F. Cukierman. Monodromy of projections. Mat. Contemp. 16, 1999.

Name and workplace of master’s thesis supervisor:

doc. Ing. Tomáš Pajdla, Ph.D., Applied Algebra and Geometry, CIIRC

Name and workplace of second master’s thesis supervisor or consultant:

Date of master’s thesis assignment: **07.01.2020** Deadline for master's thesis submission: **22.05.2020**

Assignment valid until: **30.09.2021**

doc. Ing. Tomáš Pajdla, Ph.D. Supervisor’s signature	doc. Ing. Tomáš Svoboda, Ph.D. Head of department’s signature	prof. Mgr. Petr Páta, Ph.D. Dean’s signature

III. Assignment receipt

The student acknowledges that the master’s thesis is an individual work. The student must produce his thesis without the assistance of others, with the exception of provided consultations. Within the master’s thesis, the author must state the names of consultants and include a list of references.

Date of assignment receipt	Student’s signature

Aknowledgements

I would like to thank my supervisor Tomáš Pajdla for finding a very interesting topic for research, namely symmetries in polynomial systems, and for helping me to understand the theoretical background for this topic. I would also like to thank Tim Du and Margaret Regan for providing me the computed Galois groups used in Chapter 9 and for their useful discussions about branched covers of algebraic varieties.

Declaration

I declare that the presented work was developed independently and that I have listed all sources of information used within it in accordance with the methodical instructions for observing the ethical principles in the preparation of university theses.

Prague, date ....

signature

Abstract

There are many problems in computer vision, robotics, statistics, biology, which require solving systems of polynomial equations. Every formulation of the problem by polynomial equations contains unknowns (which we are trying to determine) and parameters (which define the certain instance of the problem). For example, in computer vision, when minimal problems are formulated, the unknowns and parameters represent the camera relative poses and image measurements, respectively. Solving the problem means determining the unknowns given the parameters. However, we don't try to find a closed-form expression of unknowns as the functions of parameters since it might be very hard to do that for many problems, or because these expressions are very huge. Instead we just solve certain instances of the problem, i.e. we find the unknowns for the given values of parameters.

It may happen that the problem has symmetries. It means that there is a multivariate vector rational function such that the solution set of every instance of the problem is invariant under the action of this function. Usually, the symmetries are caused by a special formulation of the problem. For example, in computer vision, the existence of symmetries is caused by a certain geometric construction consisting of points, lines and planes: we may change this construction (i.e. the positions of points, lines and planes) without violating the relations which define this construction. In robotics, the symmetries of the inverse kinematics are caused by a special construction of the manipulator. If the problem has symmetries we can use them to simplify the problem: different solutions of every instance of the problem which are in the same orbit under the symmetry can be collapsed into one element – the solution of the instance of the reduced problem. In other words, the reduced problem with less number of solutions can be constructed from the original problem, and, as a consequence, is easier to solve.

In Chapter 9 we show how the symmetries of minimal problems in computer vision can be detected using numerical algebraic geometry. In general, the method described there can be applied to any practical problem which can be formulated by polynomial equations. After the symmetries are found, the reduced problem can be constructed. However, this is a hard task in general. We believe that in computer vision the reduced problem can be found ad hoc.

Keywords: monodromy, symmetries in polynomial systems, polynomial system simplification, symmetries in minimal problems in computer vision, Galois/monodromy group

Abstrakt

Existuje mnoho problemů v počítačovém vidění, robotice, statistice, biologii, které vyžadují řešení soustav polynomiálních rovnic. Každá formulace problému pomocí polynomiálních rovnic má neznámé (které se snážíme vypočítat) a parametry (které definují určitou instanci problému). Například, v minimálních problemech v počítačovém vidění, neznámé a parametry reprezentují relativní pozice kamer a měření z obrázků. Vyřešit problém znamená vypočítat neznáme pro zadané parametry. Nesnažíme se ale najít vzorce které vyjádří neznámé jako algebraické funkce v parametrech kvůli tomu že je to obvykle obtižné udělat nebo protože tyhle vzorce jsou obrovské. Místo toho my jen řešíme konkretní instance problému, t.j. vypočítáme neznáme pro konkretní zadané parametry.

Může se stát že problém má symetrie. To znamená že existuje nějaká racionalní funkce vůči akci které je množina řešení soustavy invariantní. Obvykle symetrie jsou způsobeny speciální formulací problému. Například, v počítačovém vidění, existence symetrií je způsobená určitou geometrickou konstrukci sestavající z bodů, přímek a rovin: tato konstrukce se dá změnit bez porušení relací které tuto konstrukci definují. V robotice, symetrie inverzní kinematické úlohy jsou způsobené speciální konstrukcí manipulátoru. Když problém má symetrie, dokážeme je použit pro zjednodušení problému: různá řešení ve stejné orbitě vůči akci symetrie se mohou zkolabovat do jednoho bodu – řešení redukovaného problému. Jinými slovy, z originálního problému dokážeme zkonstruovat redukovaný problém s menším počtem řešení který je jednodušší pro výpočet.

V 9-té kapitole ukážeme jak se dají detekovat symetrie v minimálních problemech v počítačovém vidění. Po nalezení symetrií můžeme zkonsnruovat redukovaný problém, což je obecně velmi náročný. Věříme, že v počítačovém vidění redukovaný problém se dá nalézt ad hoc.

Klíčová slova: monodromie, symetrie v polynomiálních soustavách, zjednodušení polynomiálních soustav, symetrie v minimálních problemech v počítačovém vidění, Galois/monodromy grupa

Contents

1 Introduction

1.1 Motivation

One of the most common problems in mathematics is to solve systems of polynomial equations. Nowadays, polynomial models are widely applied across the sciences. They arise in computer vision, robotics, statistics and many other branches. For example, in computer vision, we model relative camera pose problems (or minimal problems) by polynomial equations. Also, in robotics, the inverse kinematic task is modelled by polynomial equations.

Every formulation of a problem by polynomial equations contains unknowns (which we are trying to determine) and parameters (which define the certain instance of the problem). For example, in computer vision, when minimal problems are formulated, the unknowns and parameters represent the camera relative poses and image measurements, respectively. To solve the problem means to determine the unknowns given the parameters. There may exist closed-form expression of unknowns as the functions of parameters. However, we don't try to find these expressions since it might be very hard to do it for many problems, or because these expressions are very huge. Instead we just solve certain instances of the problem, i.e. we find the unknowns for the given values of parameters.

There are situations when we need to solve polynomial systems repeatedly. For example, in computer vision, when performing a 3D reconstruction from 2D images, we choose many subsets of images and for each of these subsets we solve a minimal problem associated to it. We are thus interested in speeding up the computations, since we want to make the problem solving closer to real time usage. One way to achieve this is to simplify the polynomial system by revealing its symmetries. The symmetry of the problem is a multivariate vector rational function such that the solution set of every instance of the problem is invariant under the action of this function. Usually, the symmetries are caused by a special formulation of the problem. For example, in computer vision, the existence of symmetries is caused by a certain geometric construction consisting of points, lines and planes: we may change this construction (i.e. the positions of points, lines and planes) without violating the relations which define this construction. In robotics, the symmetries of the inverse kinematics are caused by a special construction of the manipulator. If the problem has symmetries we can use them to simplify the problem: di□erent solutions of every instance of the problem which are in the same orbit under the symmetry can be collapsed into one element – the solution of the instance of the reduced problem. In other words, the reduced problem with less number of solutions can be constructed from the original problem, and, as a consequence, is easier to solve.

To verify if the polynomial system has symmetries, the so called Galois/monodromy group of this system can be computed. If a generic instance of the problem has finitely many solutions, then the Galois/monodromy group is finite and can be computed using numerical algebraic geometry. Moreover, this group encodes the structure of the solutions of a generic instance, i.e. using this group we can decide if the problem has symmetries.

1.2 State of the Art

We refer to the previous works [21][19][17]. In [21] the Galois groups of structure from motion problems were computed symbolically over the rational numbers. It is shown there that the Galois groups of the 5-point problem and the triangulation problem are the full symmetric groups S_{10} and S_6, respectively, meaning there is no structure in the solutions of generic instances of these problems. In [19] it was observed that the weak perspective-n-points problem has symmetries. They were then exploited to simplify the problem. In [17] it was shown how to find the scaling symmetries of a general polynomial system.

There is a more general approach to finding symmetries of parametric polynomial systems. For this, the so called Galois/monodromy group of the polynomial system must be computed since it encodes the structure of the solutions of this system. After that we can find the symmetries and use them for problem simplification. As far as we know, there are two general methods for Galois group computation [10][16]. The recent works [1][3] describe how to exploit the structure of the Galois/monodromy group for polynomial system simplification. However, it still remains a hard task in general.

1.3 Contributions

This work is the first attempt to develop a complete, rigorous and systematic approach to finding how to simplify parametric polynomial systems. Our motivation comes from systems appearing in computer vision but is applicable to a large family of similar structured systems which appear, e.g., in robotics [25][27] and control engineering [18].

Unlike for generic parametric polynomial systems (there are no relations among the parameters), where the situation is much simpler and has been fully characterized in [3], our situation is very complex and has not yet been described to be accessible to non-specialists.

We collected, arranged and concisely presented a large number of elements from the theory to understand the symmetries in parametric polynomial systems. Namely, we have reviewed elements from group theory, algebraic geometry, algebraic topology and Galois theory. We have shown how to combine these branches in order to understand the concept of symmetries. Our exposition covers the latest results on simplifying very generic polynomial systems [3] which have no relations among parameters. We go beyond that because problems in computer vision are more structured.

We explain the very classical and previously studied the 5-point problem using the general theory and illustrate how to analyze problems in general and suggest a possible approach to a systematic discovery of symmetries. We show that this is a very hard problem in general.

We apply our approach to two new minimal problems in computer vision discovered recently [11] and show that one can be simplified while the other cannot.

2 Elements of General Algebra and Topology

In this chapter we will give some basic facts from general algebra and topology which we will use further in Chapters 4 and 5.

2.1 General algebra

Definition 2.1. A 3-tuple $R = (R, +,)$, where each of the operations $+$, takes two elements of R and produces a new element in R, is called a commutative ring if it satisfies the following axioms:

1. $(R, +)$ is an abelian group, meaning that:

   $(a + b) + c = a + (b + c)$ for all $a, b, c \in R$ (that is, $+$ is associative).

   $a + b = b + a$ for all $a, b \in R$ (that is, $+$ is commutative).

   There is an element $0_R \in R$ such that $a + 0_R = a$ for all $a \in R$.

   For each $a \in R$ there exists $a \in R$ such that $a + (a) = 0_R$.

2. $(R,)$ is a commutative monoid, meaning that

   $(a b) c = a (b c)$ for all $a, b, c \in R$ (that is, is associative).

   $a b = b a$ for all $a, b \in R$ (that is, is commutative).

   There is an element $1_R \in R$ such that $a 1_R = 1_R a = a$ for all $a \in R$.

3. Multiplication is distributive with respect to addition, meaning that

   $a (b + c) = (a b) + (a c)$ for all $a, b, c \in R$ (left distributivity)

   $(b + c) a = (b a) + (c a)$ for all $a, b, c \in R$ (right distributivity)

It can be proved that 0_R and 1_R are unique. So, we can actually call them *the* zero element and *the* identity element, respectively.

In this thesis we work only with commutative rings and we will further write just "a ring" instead of "a commutative ring". Also, given $a, b \in R$, we will write $a b$ instead of $a + (b)$.

Definition 2.2. Let $R = (R, +,)$ be a ring. We define $I R$ to be an ideal of R, if

(i) $0_R \in I$,

(ii) $a + b \in I$ for all $a, b \in I$,

(iii) $r \cdot a \in I$ for all $r \in R$, $a \in I$.

The first natural example of an ideal is the ideal generated by a finite number of elements.

Definition 2.3. Let $a_1, \ldots, a_m$ be elements of R. Then we set

$$\langle a_1, \ldots, a_m \rangle = \left\{ \sum_{i=1}^{m} r_i \cdot a_i \;\middle|\; r_1, \ldots, r_m \in R \right\}.$$

The crucial fact is that $\langle a_1, \ldots, a_m \rangle$ is an ideal of R.

Proposition 2.4. *If $a_1, \ldots, a_m \in R$, then $\langle a_1, \ldots, a_m \rangle$ is an ideal of R. We call $\langle a_1, \ldots, a_m \rangle$ the ideal generated by $a_1, \ldots, a_m$.*

Proof. First, $0 \in \langle a_1, \ldots, a_m \rangle$ since $0 = \sum_{i=1}^{m} 0 \cdot a_i$. Next, suppose $a = \sum_{i=1}^{m} p_i \cdot a_i$ and $b = \sum_{i=1}^{m} q_i \cdot a_i$ for $p_1, \ldots, p_m, q_1, \ldots, q_m \in R$ and let $r \in R$. Then the equations

$$a + b = \sum_{i=1}^{m} (p_i + q_i) \cdot a_i,$$

$$r \cdot a = \sum_{i=1}^{m} (r \cdot p_i) \cdot a_i$$

complete the proof that $\langle a_1, \ldots, a_m \rangle$ is an ideal of R. □

Example 2.5. Let $R = (\mathbb{Z}, +, \cdot)$ and $I = n\mathbb{Z} \subset \mathbb{Z}, n \in \mathbb{Z}, n > 1$ the subset of all integers divisible by n. Then $I \subset \mathbb{Z}$ is an ideal of $\mathbb{Z}$: 0 is divisible by n; if a and b are divisible by n, so is their sum; if a is divisible by n, so is the integer $r \cdot a$ for every $r \in \mathbb{Z}$. The crucial fact is that $I = \langle n \rangle$.

Definition 2.6. Given two rings $R = (R, +_R, \cdot_R)$ and $S = (S, +_S, \cdot_S)$, a ring homomorphism from R to S is a function $\varphi\colon R \to S$ such that for all r_1 and r_2 in R it holds that

$$\varphi(r_1 +_R r_2) = \varphi(r_1) +_S \varphi(r_2),$$
$$\varphi(r_1 \cdot_R r_2) = \varphi(r_1) \cdot_S \varphi(r_2),$$
$$\varphi(1_R) = 1_S.$$

Example 2.7. Let $R = (\mathbb{Z}, +, \cdot)$ and $S = (\mathbb{Z}_n, +_{\bmod n}, \cdot_{\bmod n})$, $n \in \mathbb{N}$, $n > 1$, where $(\mathbb{Z}_n, +_{\bmod n}, \cdot_{\bmod n})$ is the ring of integers $\{0, 1, \ldots, n-1\}$ modulo n and the operations $+_{\bmod n}$ and $\cdot_{\bmod n}$ are the usual addition and multiplication of integers followed by taking modulo n. Consider the map

$$\varphi\colon \mathbb{Z} \to \mathbb{Z}_n$$
$$a \mapsto a \bmod n$$

It is a ring homomorphism since

$$\varphi(a+b) = (a+b) \bmod n = (a \bmod n) +_{\text{mod } n} (b \bmod n) = \varphi(a) +_{\text{mod } n} \varphi(b),$$

$$\varphi(a \quad b) = (a \quad b) \bmod n = (a \bmod n) \quad _{\text{mod } n} (b \bmod n) = \varphi(a) \quad _{\text{mod } n} \varphi(b),$$

$$\varphi(1_{\mathbb{Z}}) = 1 \bmod n = 1_{\mathbb{Z}_n}.$$

A bijective ring homomorphism has a special name.

Definition 2.8. Given two rings $R = (R, +_R, \ _R)$ and $S = (S, +_S, \ _S)$, a ring isomorphism from R to S is a bijective ring homomorphism from R to S. Then rings R and S are said to be isomorphic.

For simplicity, we will further write just R for a ring $R = (R, +, \ )$. And for two rings R, S we write $+$ (resp. $\ $) for both $+_R$ and $+_S$ (resp. $\ _R$ and $\ _S$). Also in some cases, for simplicity, we will omit the sign $\ $ in $a \ b$ and write just ab for $a, b \in R$.

Proposition 2.9. *Any ring homomorphism $\varphi\colon R \to S$ sends 0_R to 0_S. Moreover, $\varphi(\ r) = \ \varphi(r)$ for all $r \in R$.*

Proof. We can write:

$$\varphi(0_R) = \varphi(0_R + 0_R) = \varphi(0_R) + \varphi(0_R).$$

Then

$$\varphi(0_R) + 0_S = \varphi(0_R) + \varphi(0_R) + 0_S = \varphi(0_R) + \varphi(0_R) \ \Rightarrow \ 0_S = \varphi(0_R).$$

For the moreover part:

$$0_S = \varphi(0_R) = \varphi(r \quad r) = \varphi(r) + \varphi(\ r) \ \Rightarrow \ \varphi(\ r) = \quad \varphi(r).$$

□

Definition 2.10. Let $\varphi\colon R \to S$ be a ring homomorphism. We define the kernel of φ to be

$$\ker(\varphi) = \{r \in R : \varphi(r) = 0_S\}$$

and the image of φ to be

$$\text{im}(\varphi) = \{\varphi(r) : r \in R\}.$$

Example 2.11. Take the ring homomorphism φ from Example 2.7. Its kernel $\ker(\varphi)$ is the set of integers which are divisible by n, so $\ker(\varphi) = n\mathbb{Z}$. The image $\text{im}(\varphi) = \mathbb{Z}_n$, because the integers $0, 1, ..., n \quad 1 \in \mathbb{Z}$ map to $0, 1, ..., n \quad 1 \in \mathbb{Z}_n$ by φ, respectively.

Given a ring R and an ideal $I \quad R$, we define an equivalence relation $\quad$ on R as follows:

$$a \quad b \quad () \quad a \quad b \in I \tag{2.1}$$

It is easy to verify that this is indeed an equivalence relation: $a \quad a$ since $0_R \in I$, $a \quad b$ implies $b \quad a$ since $b \quad a = (\ 1) \ (a \quad b) \in I$, $a \quad b$ and $b \quad c$ imply $a \quad c$ since $(a \quad b) + (b \quad c) = a \quad c \in I$. We can then define the equivalence class of $a \in R$ as

$$[a] \stackrel{\text{def}}{=} \{b \in R \mid a \quad b\}.$$

It follows from (2.1) that

$$[a] = a + I = \{a + r \mid r \in I\}.$$

We call $[a]$ a coset of I in R given by a. The set of all such equivalence classes is denoted by R/I. It becomes a ring, the quotient ring of R modulo I, if one defines

$$(a + I) + (b + I) \stackrel{\text{def}}{=} (a + b) + I,$$

$$(a + I) \cdot (b + I) \stackrel{\text{def}}{=} (a \cdot b) + I.$$

The zero and the identity elements are $0_R + I$ and $1_R + I$, respectively. The following proposition shows that these operations are well-defined, meaning the result of the sum and the product doesn't depend on the choice of class representative.

Proposition 2.12. *The operations above, which turn R/I into a ring, are well-defined.*

Proof. Notice that it is enough to check it only for two different representatives of $a+I$ since R is commutative. So, take $a_1 + I = a_2 + I$ and $b+I$. Then $a_1 = a_2 + j$ for some $j \in I$. We have

$$(a_1 + I) + (b + I) = (a_1 + b) + I = (a_2 + j + b) + I = (a_2 + b) + I = (a_2 + I) + (b + I),$$

$$(a_1 + I) \cdot (b + I) = (a_1 \cdot b) + I = ((a_2 + j) \cdot b) + I = (a_2 \cdot b + j \cdot b) + I =$$
$$= (a_2 \cdot b) + I = (a_2 + I) \cdot (b + I).$$

□

Example 2.13. Recall Example 2.5. The quotient ring $R/I = \mathbb{Z}/n\mathbb{Z}$ consists of the equivalence classes $[a] = a + n\mathbb{Z}$. Each of the equivalence classes $[a]$ is the set of integers congruent to a modulo n.

We now give the First Isomorphism Theorem for rings.

Proposition 2.14. *Let $\varphi\colon R \to S$ be a ring homomorphism. Then,*

1. $\ker(\varphi)$ *is an ideal of* R.
2. $\mathrm{im}(\varphi)$ *is a subring of* S.
3. *The quotient ring* $R/\ker(\varphi)$ *is isomorphic to* $\mathrm{im}(\varphi)$.

In particular, if φ is surjective then S is isomorphic to $R/\ker(\varphi)$.

Proof. 1. Take $k_1, k_2 \in \ker(\varphi)$ and $r \in R$. Then:

$$\varphi(k_1 + k_2) = \varphi(k_1) + \varphi(k_2) = 0_S + 0_S = 0_S \Rightarrow k_1 + k_2 \in \ker(\varphi).$$

$$\varphi(r \cdot k_1) = \varphi(r) \cdot \varphi(k_1) = \varphi(r) \cdot 0_S = 0_S \Rightarrow r \cdot k_1 \in \ker(\varphi).$$

2. According to Definition 2.1, we need to check 4 things:

The set $\mathrm{im}(\varphi)$ is closed under the operations $+$ and $\cdot$. To check this we take $s_1 = \varphi(r_1), s_2 = \varphi(r_2)$ 2 $\mathrm{im}(\varphi)$. Then

$$s_1 + s_2 = \varphi(r_1) + \varphi(r_2) = \varphi(r_1 + r_2) \text{ 2 } \mathrm{im}(\varphi),$$

$$s_1 \cdot s_2 = \varphi(r_1) \cdot \varphi(r_2) = \varphi(r_1 \cdot r_2) \text{ 2 } \mathrm{im}(\varphi).$$

By Proposition 2.9, $0_S = \varphi(0_R)$ 2 $\mathrm{im}(\varphi)$.

By Proposition 2.9, for any element $s = \varphi(r)$ 2 $\mathrm{im}(\varphi)$ there is $-s = \varphi(-r)$ 2 $\mathrm{im}(\varphi)$.

By Definition 2.6, $1_S = \varphi(1_R)$ 2 $\mathrm{im}(\varphi)$.

3. Let $I = \ker(\varphi)$. Define a map

$$\overline{\varphi}\colon R/I \to \mathrm{im}(\varphi)$$
$$r + I \mapsto \varphi(r)$$

It is well-defined because if $r_1 + I = r_2 + I$, meaning $r_1 = r_2 + j$ for some j 2 I, then

$$\overline{\varphi}(r_1 + I) = \varphi(r_1) = \varphi(r_2 + j) = \varphi(r_2) + \varphi(j) = \varphi(r_2) + 0_S = \varphi(r_2) = \overline{\varphi}(r_2 + I).$$

It is a ring homomorphism because

$$\overline{\varphi}((r_1+I)+(r_2+I)) = \overline{\varphi}((r_1+r_2)+I) = \varphi(r_1+r_2) = \varphi(r_1)+\varphi(r_2) = \overline{\varphi}(r_1+I)+\overline{\varphi}(r_2+I),$$

$$\overline{\varphi}((r_1 + I) \cdot (r_2 + I)) = \overline{\varphi}((r_1 \cdot r_2) + I) = \varphi(r_1 \cdot r_2) = \varphi(r_1) \cdot \varphi(r_2) = \overline{\varphi}(r_1 + I) \cdot \overline{\varphi}(r_2 + I),$$

$$\overline{\varphi}(1_R + I) = \varphi(1_R) = 1_S.$$

The map $\overline{\varphi}$ is injective because

$$\overline{\varphi}(r_1 + I) = \overline{\varphi}(r_2 + I) \Rightarrow \varphi(r_1) = \varphi(r_2) \Rightarrow \varphi(r_1 - r_2) = 0_S \Rightarrow r_1 - r_2 \text{ 2 } I \Rightarrow r_1 + I = r_2 + I.$$

It is obviously surjective because for any $s = \varphi(r)$ 2 $\mathrm{im}(\varphi)$ there is a coset $r + I$ which maps to s by $\overline{\varphi}$. □

Despite the fact that the following proposition is a corollary of Proposition 2.14, we prove it in another way.

Proposition 2.15. *A ring homomorphism $\varphi\colon R \to S$ is injective if and only if $\ker(\varphi) = \{0_R\}$.*

Proof. If φ is injective then obviously there is only one element which maps to 0_S by φ, namely 0_R. Conversely, suppose $\ker(\varphi) = \{0_R\}$. Then

$$\varphi(r_1) = \varphi(r_2) \Rightarrow \varphi(r_1 - r_2) = 0_S \Rightarrow r_1 - r_2 = 0_R \Rightarrow r_1 = r_2.$$

□

Remark 2.16. From now on, if we have a surjective ring homomorphism $\varphi\colon R \to S$ with kernel $I \subsetneq R$, the induced isomorphism between the quotient ring R/I and S will be denoted as $\overline{\varphi}\colon R/I \to S$ and we will write

$$R/I \overset{\overline{\varphi}}{=} S.$$

By $\varphi\colon R \hookrightarrow S$ we denote an injective ring homomorphism.

Example 2.17. Consider the ring homomorphism from Example 2.7:

$$\begin{aligned} \varphi\colon \mathbb{Z} &\to \mathbb{Z}_n \\ a &\mapsto a \bmod n \end{aligned}$$

As it is surjective and its kernel is $n\mathbb{Z}$, by Proposition 2.14, we have

$$\mathbb{Z}/n\mathbb{Z} \overset{\overline{\varphi}}{=} \mathbb{Z}_n.$$

Definition 2.18. A nonzero ring R (i.e. $R \neq \{0_R\}$) is called an integral domain if it has no nontrivial zero divisors. In other words, if $a, b \in R$, then

$$ab = 0_R \implies a = 0_R \text{ or } b = 0_R.$$

Example 2.19. The ring of integer numbers is an integral domain since there are no two integers $a, b \neq 0$ such that $ab = 0$. The ring of integers modulo 4 is not an integral domain because $2 \cdot 2 = 0$ in $\mathbb{Z}_4$.

Definition 2.20. A field F is a nonzero ring where every nonzero element has a multiplicative inverse, meaning, to every $a \in F, a \neq 0_F$ there is $b \in F$ such that $ab = 1_F$.

Example 2.21. The set $\mathbb{Q}$ of rational numbers is a field. To every fraction $\frac{a}{b}$ with $a \neq 0$ there is a multiplicative inverse $\frac{b}{a}$.

It is good to know how the rational numbers $\mathbb{Q}$ are actually constructed. We can take the ring of integers $\mathbb{Z}$ and consider the set $\mathbb{Z} \times \mathbb{Z}^*$, where $\mathbb{Z}^* = \mathbb{Z} \setminus \{0\}$. We can define a relation on this set as

$$(a, b) \sim (c, d) \iff ad = bc \tag{2.2}$$

It can be verified that this is an equivalence relation. Then we can form the set of equivalence classes $(\mathbb{Z} \times \mathbb{Z}^*)/\sim$. We put a ring structure on this set as follows:

$$[(a, b)] + [(c, d)] \overset{\text{def}}{=} [(ad + bc, bd)], \quad [(a, b)] \cdot [(c, d)] \overset{\text{def}}{=} [(ac, bd)] \tag{2.3}$$

It can be verified that these operations are well-defined. Finally, it can be checked that $(\mathbb{Z} \times \mathbb{Z}^*)/\sim$ is actually a field. We define

$$\mathbb{Q} \overset{\text{def}}{=} (\mathbb{Z} \times \mathbb{Z}^*)/\sim .$$

We generalize the above construction of $\mathbb{Q}$ as follows.

Proposition 2.22. *Let R be an integral domain. If we define a relation on $R \quad R$ as in (2.2) and put a ring structure on it as in (2.3), then $(R \quad R\ )/\quad$ becomes a field. We call it the field of fractions of R and denote it $\mathrm{Frac}(R)$.*

Proof. 1. We at first verify that (2.2) is indeed an equivalence relation. We need to prove that it is reflexive, symmetric and transitive. For reflexivity, take (a,b) *2* $R \quad R$. Then $ab = ba$ since R is commutative. This means $(a,b) \quad (a,b)$. For symmetry, take $(a,b),(c,d)$ *2* $R \quad R$ and suppose $(a,b) \quad (c,d)$. Then $ad = bc$. But then $cb = da$, since R is commutative, and so $(c,d) \quad (a,b)$. For transitivity, take $(a,b),(c,d),(e,f)$ *2* $R \quad R$ and suppose $(a,b) \quad (c,d),(c,d) \quad (e,f)$. Then

$$(af)d = (ad)f = (bc)f = b(cf) = (be)d.$$

As (a,d) *2* $R \quad R$, then d *6*$= 0_R$. Hence $(af)d = (be)d$ implies $(af \quad be)d = 0_R$. Since R is an integral domain, this implies $af \quad be = 0_R$, or that $af = be$. Thus, $(a,b) \quad (e,f)$.

2. Now we prove that the operations (2.3) are well-defined. Suppose that $[(a,b)] = [(a',b')]$ and $[(c,d)] = [(c',d')]$. Then $ab' = a'b$ and $cd' = c'd$. Thus,

$$(ad+bc)b'd' = adb'd' + bcb'd' = a'bdd' + c'dbb' = (a'd' + b'c')bd,$$

$$acb'd' = a'bc'd = bda'c',$$

which means $[(ad+bc,bd)] = [(a'd' + b'c', b'd')]$ and $[(ac,bd)] = [(a'c', b'd')]$.

3. It is straightforward to verify that $\mathrm{Frac}(R)$ forms a ring under these operations. The zero element is $0_{\mathrm{Frac}(R)} = [(0_R, 1_R)]$ and the identity element is $1_{\mathrm{Frac}(R)} = [(1_R, 1_R)]$. Given an element $[(a,b)]$ *2* $\mathrm{Frac}(R)$ with a *6*$= 0_R$ we can see that $[(b,a)]$ is the multiplicative inverse of $[(a,b)]$. It follows that $\mathrm{Frac}(R)$ is a field.

□

Remark 2.23. Since for any ring R its field of fractions is defined in the same way as the rational numbers Q were defined, every element $[(a,b)]$ of $\mathrm{Frac}(R)$ will be denoted $\left[\frac{a}{b}\right]$.

We can embed R to $\mathrm{Frac}(R)$ in the same way we embed Z into Q. We do it via the following ring homomorphism:

$$\psi\colon R \ ! \ \mathrm{Frac}(R)$$
$$a \ 7! \ \left[\frac{a}{1_R}\right]$$

We can see that ψ is injective since

$$\left[\frac{a}{1_R}\right] = \left[\frac{b}{1_R}\right] \ =) \quad a = a1_R = b1_R = b.$$

Proposition 2.24. *Let $\varphi\colon R \ ! \ F$ be an injective ring homomorphism where R is an integral domain and F is a field. Then there exists an injective ring homomorphism $\varphi : \mathrm{Frac}(R) \ ! \ F$ which extends φ, meaning $\varphi\ \big|_{\psi(R)} = \varphi$.*

Proof. We define

$$\varphi : \mathrm{Frac}(R) \to F$$
$$\left[\frac{a}{b}\right] \mapsto \varphi(a)\varphi(b)^{-1}$$

1. We check that it is well-defined. At first notice that $\varphi(b)$ has an inverse in F since it is nonzero for every nonzero b (follows from injectivity of φ). Now, let $\frac{a}{b} \sim \frac{a'}{b'}$ which means $ab' = a'b$. Then from

$$\varphi(a')\varphi(b) = \varphi(a'b) = \varphi(ab') = \varphi(a)\varphi(b') \Rightarrow \varphi(a')\varphi(b')^{-1} = \varphi(a)\varphi(b)^{-1}$$

   it follows that

$$\varphi\left(\left[\frac{a'}{b'}\right]\right) = \varphi(a')\varphi(b')^{-1} = \varphi(a)\varphi(b)^{-1} = \varphi\left(\left[\frac{a}{b}\right]\right).$$

2. φ is a ring homomorphism since

$$\varphi\left(\left[\frac{a}{b}\right] + \left[\frac{c}{d}\right]\right) = \varphi\left(\left[\frac{ad+bc}{bd}\right]\right) = \varphi(ad+bc)\varphi(bd)^{-1} =$$
$$= (\varphi(a)\varphi(d) + \varphi(b)\varphi(c))\underbrace{(\varphi(b)\varphi(d))^{-1}}_{\varphi(d)^{-1}\varphi(b)^{-1}} = \varphi(a)\varphi(b)^{-1} + \varphi(c)\varphi(d)^{-1} =$$
$$= \varphi\left(\left[\frac{a}{b}\right]\right) + \varphi\left(\left[\frac{c}{d}\right]\right),$$
$$\varphi\left(\left[\frac{a}{b}\right]\left[\frac{c}{d}\right]\right) = \varphi\left(\frac{ac}{bd}\right) = \varphi(ac)\varphi(bd)^{-1} = \varphi(a)\varphi(b)^{-1}\varphi(c)\varphi(d)^{-1} =$$
$$= \varphi\left(\left[\frac{a}{b}\right]\right)\varphi\left(\left[\frac{c}{d}\right]\right),$$
$$\varphi\left(\left[\frac{1_R}{1_R}\right]\right) = \varphi(1_R)\varphi(1_R)^{-1} = 1_F 1_F^{-1} = 1_F.$$

3. To prove $\varphi\big|_{\psi(R)} = \varphi$ we just notice that

$$\varphi\left(\left[\frac{a}{1_R}\right]\right) = \varphi(a)\varphi(1_R)^{-1} = \varphi(a)1_R^{-1} = \varphi(a).$$

□

Definition 2.25. Let R be a ring and $I \subsetneq R$ be an ideal of R. Then I is said to be prime, if, for $a, b \in R$, we have

$$ab \in I \Rightarrow a \in I \text{ or } b \in I.$$

Example 2.26. Let $R = \mathbb{Z}$ and $I = p\mathbb{Z} \subsetneq \mathbb{Z}$ for some prime number p. To see that I is a prime ideal let $a, b \in \mathbb{Z}$ and $ab \in I$. Then p divides ab. Because p is prime, then p divides a or p divides b. This exactly means that $a \in I$ or $b \in I$.

Proposition 2.27. *Let R be a ring and I R be an ideal of R. Then R/I is an integral domain if and only if I is prime.*

Proof. Let a, b 2 R. For R/I to be an integral domain means the following:

$$ab + I = (a + I)(b + I) = 0_R + I \ =) \quad a + I = 0_R + I \text{ or } b + I = 0_R + I.$$

By (2.1), this is equivalent to:

$$ab \ 2 \ I \ =) \quad a \ 2 \ I \text{ or } b \ 2 \ I.$$

□

2.2 General topology

It wasn't our aim here to write an introduction to the general topology. This section looks like a list of definitions and propositions to which we will refer in Chapters 4 and 5. So, you can actually skip this section and return back when you see a reference to it in Chapters 4 and 5.

Definition 2.28. A topological space is an ordered pair (X, τ), where X is a set and τ is a collection of subsets of X, satisfying the following axioms:

1. ? and X belong to τ.
2. Arbitrary (finite or infinite) intersections of elements from τ belong to τ.
3. Finite unions of elements from τ belong to τ.

The elements of τ are called closed subsets of X and the collection τ is called a topology on X. A subset Y X is said to be open in X if its complement $X \ n \ Y$ is closed in X, i.e. $X \ n \ Y \ 2 \ \tau$.

Proposition 2.29. *Let (X, τ) be a topological space. Then:*

1. *Arbitrary (finite or infinite) unions of open subsets of X are open in X.*
2. *Finite intersections of open subsets of X are open in X.*

Proof. 1. Let $fU_\alpha g_{\alpha 2 A}$ be an arbitrary collection of open subsets of X. Thus every U_α is of the form $X \ n \ V_\alpha$ for a closed subset V_α X. Then using De Morgan's law:

$$\bigcup_{\alpha 2 A} U_\alpha = \bigcup_{\alpha 2 A} (X \ n \ V_\alpha) = X \setminus \left(\bigcap_{\alpha 2 A} V_\alpha \right).$$

Since $\bigcap_{\alpha 2 A} V_\alpha$ is closed in X, then $X \setminus \left(\bigcap_{\alpha 2 A} V_\alpha \right)$ is open in X.

2. Let $\{U_i\}_{i=1}^n$ be a finite collection of open sets of X. Thus, every U_i is of the form $X \setminus V_i$ for a closed subset $V_i \subset X$. Using again De Morgan's law we obtain:

$$\bigcap_{i=1}^{n} U_i = \bigcap_{i=1}^{n} (X \setminus V_i) = X \setminus \left(\bigcup_{i=1}^{n} V_i \right).$$

Since $\bigcup_{i=1}^{n} V_i$ is closed in X, then $X \setminus \left(\bigcup_{i=1}^{n} V_i \right)$ is open in X.

□

Proposition 2.30. *Let (X, τ) be a topological space and $Y \subset X$ be a subset. Let*

$$\tau' = \{V \cap Y \mid V \in \tau\}.$$

Then (Y, τ') is a topological space. The topology τ' is called the subspace topology on Y.

Proof. 1. $\emptyset$ and Y belong to τ' since

$$\emptyset = \emptyset \cap Y, \quad Y = X \cap Y.$$

2. Let $\{V'_\alpha\}_{\alpha \in A}$ be an arbitrary collection of elements from τ'. Then each of them is of the form $V'_\alpha = V_\alpha \cap Y$ for some $V_\alpha \in \tau$. Thus,

$$\bigcap_{\alpha \in A} V'_\alpha = \bigcap_{\alpha \in A} \left(V_\alpha \cap Y \right) = \left(\bigcap_{\alpha \in A} V_\alpha \right) \cap Y \in \tau'.$$

3. Let $\{V'_i\}_{i=1}^n$ be a finite collection of elements from τ'. Then $V'_i = V_i \cap Y$ for some $V_i \in \tau$ and hence

$$\bigcup_{i=1}^{n} V'_i = \bigcup_{i=1}^{n} \left(V_i \cap Y \right) = \left(\bigcup_{i=1}^{n} V_i \right) \cap Y \in \tau'.$$

□

Definition 2.31. Let (X, τ) be a topological space and $Y \subset X$ be a subset. The closure $\overline{Y}$ of Y in X is the intersection of all closed subsets of X which contain Y. In other words,

$$\overline{Y} = \bigcap_{Y \subset V \subset X}^{\text{closed}} V.$$

Since an arbitrary intersection of closed sets is closed, then $\overline{Y}$ is closed in X.

Corollary 2.32. *Let (X, τ) be a topological space and let $Y \subset X$ be a closed subset. Then for any subset $S \subset Y$ there holds*

$$\overline{S} \subset Y,$$

where $\overline{S}$ denotes the closure of S in X.

Proposition 2.33. *Let (X, τ) be a topological space and let $Y \quad X$. If for any y 2 Y there exists an open subset U_y of X which contains y and is contained in Y, then Y is open in X.*

Proof. We can write

$$Y = \bigcup_{y2Y} U_y.$$

Because U_y is open in X, then using Proposition 2.29 we conclude that Y is open in X. □

Definition 2.34. Given a topological space (X, τ) and a point p 2 X, a neighbourhood of p in X is a subset $V \quad X$ such that contains an open subset U containing p,

$$p \; 2 \; U \quad V.$$

If V is an open subset of X, then V is called an open neighbourhood of p.

Definition 2.35. Let (X, τ_1) and (Y, τ_2) be two topological spaces. Then a function $f : X \; ! \; Y$ is said to be continuous if for every open subset $U \quad Y$, the inverse $f \; {}^1(U)$ is an open subset of X.

Definition 2.36. A homeomorphism is a continuous map between topological spaces which is bijective and has a continuous inverse.

Definition 2.37. A topological space (X, τ) is connected if any presentation of X as $X = U_1 \; [\; U_2$ by disjoint open subsets implies $U_1 = X$ or $U_2 = X$.

Definition 2.38. Let (X, τ) be a topological space and x_1, x_2 2 X. A path in X from x_1 to x_2 is a continuous function from the unit interval $I = [0, 1] \quad$ R to X such that $f(0) = x_1$ and $f(1) = x_2$.

Definition 2.39. A topological space (X, τ) is path-connected if for any x_1, x_2 2 X there is a path in X from x_1 to x_2.

Definition 2.40. A topological space (X, τ) is locally path-connected if for any x 2 X and any open subset V_x 3 x of X there is a smaller open subset x 2 $U_x \quad V_x$ of X which is path-connected in the subspace topology.

Proposition 2.41. *Let (X, τ) be a topological space. If X is connected and locally path-connected, then X is path-connected.*

Proof. Fix x 2 X and let Y be the set of all points in X such that there is a path from x to any point in Y. The set Y is nonempty, since X is locally path-connected. We would like to show that Y is both open and closed in X. Let $Y^c = XnY$. Then we can represent X

$$X = Y \; [\; Y^c$$

as a union of disjoint open subsets of X. By Definition 2.37, we have $X = Y$, since Y is nonempty. Since Y is path-connected, then so is X.

To show that Y is open in X, let y 2 Y. Since X is locally path-connected, we can choose an open subset U 3 y of X which is path-connected. Thus, for any u 2 U there is a

path from u to y and, since $y \in Y$, there is a path from y to x. Hence there is a path from u to x in X. This means that $U \subseteq Y$. By Proposition 2.33, Y is open in X.

To show that Y is closed in X we will prove that Y^c is open in X. Let $y' \in Y^c$ and choose an open subset $U' \ni y'$ of X which is path-connected. The intersection $U' \setminus Y = \emptyset$, since if there is some $p \in U' \setminus Y$, then there is a path in X from y' to p ($p \in U'$) and also there is a path in X from p to x ($p \in Y$). But this means that there is a path in X from y' to x, so that $y' \in Y$. It's a contradiction. Thus, $U' \setminus Y = \emptyset$ and then $U' \subseteq Y^c$. Again, by Proposition 2.33, Y^c is open in X. □

Definition 2.42. A topological space (X, τ) is said to be reducible if it can be written as a union $X = X_1 \cup X_2$ of two proper closed subsets X_1, X_2 of X. A topological space is irreducible if it is not reducible.

Definition 2.43. Let (X, τ) be an irreducible topological space. A subset $Y \subseteq X$ is said to be dense in X if $\overline{Y} = X$.

Proposition 2.44. *Let (X, τ) be an irreducible topological space and let $Y \subseteq X$ be a nonempty open subset. Then Y is dense in X.*

Proof. For contradiction suppose $\overline{Y} \neq X$. Since Y is a nonempty open subset of X, then $Y^c = X \setminus Y$ is a proper closed subset of X. Since $Y \subseteq \overline{Y}$, then

$$X = Y^c \cup \overline{Y}.$$

So, we are able to write X as a union of two proper closed subsets of X, which means X is reducible. It is a contradiction. □

Proposition 2.45. *Let (X, τ) be a topological space. If there is a finite collection $\{X_i\}_{i=1}^n$ of proper closed subsets of X such that $X = \bigcup_{i=1}^{n} X_i$, then X is reducible.*

Proof. From all possible finite covers of X by proper closed subsets (which is nonempty since $\{X_i\}_{i=1}^n$ is such a cover) take the one with the smallest number m of subsets. Denote it as $\{Y_i\}_{i=1}^m$. We know that

$$X = \bigcup_{i=1}^{m} Y_i = \left(\bigcup_{i=1}^{m-2} Y_i \right) \cup \left(Y_{m-1} \cup Y_m \right).$$

By minimality of m it follows that $Y_{m-1} \cup Y_m$ is not a proper subset, because otherwise we would get a cover by $m-1$ proper closed subsets of X. Thus $Y_{m-1} \cup Y_m = X$. So, X is reducible.

□

Proposition 2.46. *Let (X, τ) be an irreducible topological space. Then the intersection of a finite number of nonempty open sets of X is nonempty and open in X.*

Proof. Denote a finite collection of nonempty finite open sets of X as $fU_ig_{i=1}^n$. According to Definition 2.29, their intersection $\bigcap_{i=1}^{n} U_i$ is open. We can write $U_i = X \setminus Y_i$ for some proper closed subset Y_i of X. For contradiction suppose $\bigcap_{i=1}^{n} U_i = ?$. Then

$$? = \bigcap_{i=1}^{n} U_i = \bigcap_{i=1}^{n} \left(X \setminus Y_i \right) = X \setminus \bigcup_{i=1}^{n} Y_i \quad) \quad \bigcup_{i=1}^{n} Y_i = X.$$

So, X can be written as a union of proper closed subsets of X. By Proposition 2.45, X is reducible. It is a contradiction. □

3 Elements of Group Theory

Group theory is crucial for understanding the main ideas of this work. The word "group" was invented by a french matematician Évariste Galois in 19th century who used this object to study permutation of the roots of a univariate polynomial. In some sense we will use groups in this work for the same purpose (you will learn more about it in Chapter 7). Here we are going to explain basic definitions of group theory (group, group homomorphism, etc.) and, finally, explain the relation between the stabilizer, normalizer and centralizer. We also would like to note that there is a powerful software for computation with groups, called GAP [12], which we used in this work.

3.1 Basic definitions

Definition 3.1. A tuple $G = (G, \ )$, where the operation $\ $ takes two elements and produces another element, denoted $a \ \ b$, is called a group if it satisfies the following axioms:

(i) Closure: For all a, b in G, the result of the operation, $a \ \ b$, is also in G.

(ii) Associativity: For all a,b and c in G, $(a \ \ b) \ \ c = a \ \ (b \ \ c)$.

(iii) Identity element: There exists an element 1_G in G such that, for every element a in G, the equation $1_G \ \ a = a \ \ 1_G = a$ holds true. It can be proved that such an element is unique, and thus one speaks of *the* identity element.

(iv) Inverse element: For each a in G, there exists an element $a^{\ 1}$ in G, such that $a \ \ a^{\ 1} = a^{\ 1} \ \ a = 1_G$.

Example 3.2. The set of integers Z together with the addition operation $+$ forms a group $(\mathbb{Z}, +)$. The sum of two integers is obviously an integer (closure). The addition operation is associative. The identity element is $0 \in \mathbb{Z}$. The inverse element to $a \in \mathbb{Z}$ is $\ a$.

Definition 3.3. We say that a group $G = (G, \ )$ is finite if G is a finite set. The order of G is the number of elements of G.

Definition 3.4. Given a group $(G, \ )$ and a subset $H \ \ G$, a tuple $(H, \ )$ is called a subgroup of $(G, \ )$ if $(H, \ )$ is a group.

Example 3.5. Let $(G, \ ) = (\mathbb{Z}, +)$ and $H = n\mathbb{Z} \ \ \mathbb{Z}$, where $n\mathbb{Z}$ is the set of integers that are divisible by n. We verify that $(n\mathbb{Z}, +)$ is a subgroup of $(\mathbb{Z}, +)$. The sum of two integers divisible by n is again an integer divisible by n (closure). Operation $+$ is associative in $n\mathbb{Z}$,

because it is associative in Z. The identity element in $n\mathbb{Z}$ is 0. And to every integer divisible by n there is an integer with an opposite sign, which is obviously divisible by n.

Definition 3.6. Given two groups $(G, \ )$ and $(H, \ )$, a group homomorphism from $(G, \ )$ to $(H, \ )$ is a function $\varphi : G \ ! \ H$ such that for all g_1 and g_2 in G it holds that

$$\varphi(g_1 \ \ g_2) = \varphi(g_1) \ \ \varphi(g_2).$$

Example 3.7. Consider $(G, \ ) = (\mathbb{Z}, +)$ and $(H, \ ) = (\mathbb{Z}_n, +_{\bmod n})$, $n \ 2 \ \mathbb{Z}, n > 1$, where $(\mathbb{Z}_n, +_{\bmod n})$ is the group of integers $f0, 1, \ldots, n \ \ 1g$ modulo n and the operation $+_{\bmod n}$ is the usual addition followed by taking modulo n. Define a map

$$\begin{aligned} \varphi\colon \mathbb{Z} \ &! \ \mathbb{Z}_n \\ a \ &7! \ a \bmod n \end{aligned}$$

It is easy to verify that φ is a group homomorphism, because

$$\varphi(a+b) = (a+b) \bmod n = (a \bmod n) +_{\bmod n} (b \bmod n) = \varphi(a) +_{\bmod n} \varphi(b).$$

Proposition 3.8. *Any group homomorphism $\varphi\colon G \ ! \ H$ sends 1_G to 1_H. Moreover, $\varphi(g \ ^{1}) = (\varphi(g)) \ ^{1}$ for all $g \ 2 \ G$.*

Proof. By Definition 3.6,

$$\varphi(g_1 \ \ g_2) = \varphi(g_1) \ \ \varphi(g_2) \quad 8g_1, g_2 \ 2 \ G.$$

Take $g_1 = g_2 = 1_G$. Then

$$\varphi(1_G) = \varphi(1_G \ \ 1_G) = \varphi(1_G) \ \ \varphi(1_G).$$

Multiplying by $(\varphi(1_G)) \ ^{1}$ from the both sides we obtain

$$1_H = \varphi(1_G).$$

For the last part of the statement:

$$1_H = \varphi(1_G) = \varphi(gg \ ^{1}) = \varphi(g)\varphi(g \ ^{1}) \ ) \ \varphi(g \ ^{1}) = (\varphi(g)) \ ^{1}.$$

□

Definition 3.9. Given two groups $(G, \ )$ and $(H, \ )$, a group isomorphism from $(G, \ )$ to $(H, \ )$ is a bijective group homomorphism φ from $(G, \ )$ to $(H, \ )$. Then groups $(G, \ )$ and $(H, \ )$ are said to be isomorphic and we write

$$G \overset{\varphi}{=} H.$$

Example 3.10. Let $(G, \ ) = (\mathbb{Z}, +)$ and $(H, \ ) = (n\mathbb{Z}, +)$ for some $n \ 2 \ \mathbb{Z}$, $n > 1$. Then define φ to be:

$$\begin{aligned} \varphi\colon \mathbb{Z} \ &! \ n\mathbb{Z} \\ a \ &7! \ n \ \ a \end{aligned}$$

where is the usual multiplication of integers. We claim that φ is a group homomorphism since

$$\varphi(a+b) = n \quad (a+b) = n \quad a + n \quad b = \varphi(a) + \varphi(b).$$

We can also show that φ is bijective. Its inverse is

$$\begin{aligned} \varphi^{-1} &: n\mathbb{Z} \to \mathbb{Z} \\ a &\mapsto \frac{1}{n} \quad a \end{aligned}$$

This shows that φ is an isomorphism and that the groups $\mathbb{Z}$ and $n\mathbb{Z}$ are isomorphic.

Further in the text we will omit the group operation and write just G for a group $(G, \ )$.

Definition 3.11. Let $\varphi : G \to H$ be a group homomorphism. We define the kernel of φ to be

$$\ker(\varphi) \stackrel{\text{def}}{=} \{g \in G : \varphi(g) = 1_H\}$$

and the image of φ to be

$$\operatorname{im}(\varphi) \stackrel{\text{def}}{=} \{\varphi(g) : g \in G\}.$$

Example 3.12. Let $G = \mathbb{Z}$ and $H = \mathbb{Z}_n, n \in \mathbb{Z}, n > 1$. We define φ to be

$$\begin{aligned} \varphi &: \mathbb{Z} \to \mathbb{Z}_n \\ a &\mapsto a \bmod n \end{aligned}$$

Then the kernel of φ is exactly the set of integers which give zero modulo n. These are exactly the integers from $n\mathbb{Z}$. So, $\ker(\varphi) = n\mathbb{Z}$. Here $\operatorname{im}(\varphi) = \mathbb{Z}_n$, because the integers $0, 1, ..., n \quad 1 \in \mathbb{Z}$ map to $0, 1, ..., n \quad 1 \in \mathbb{Z}_n$ by φ, respectively.

Definition 3.13. Given an element g of a group G and a subgroup H of G, define

$$gH = \{gh \mid h \in H\}$$

to be the left coset of H in G with respect to g.

Example 3.14. Let $G = \mathbb{Z}$ and $H = n\mathbb{Z}$. Take any $g \in \mathbb{Z}$. Then

$$gH = \{g + nk \mid k \in \mathbb{Z}\}$$

is the set of integers congruent to g modulo n.

Proposition 3.15. *Let G be a group and H be a subgroup of G. Let g_1H and g_2H be two left cosets of H in G. Then either $g_1H \cap g_2H = \emptyset$ or $g_1H = g_2H$.*

Proof. If $g_1H \cap g_2H = \emptyset$, then we are done. Suppose $g_1H \cap g_2H \neq \emptyset$. Take $a \in g_1H \cap g_2H$. Then $a = g_1h_1 = g_2h_2$ for some $h_1, h_2 \in H$. This means $g_2 = g_1h_1h_2^{-1}$, or that $g_2 = g_1h$ for some $h \in H$. It is easy to see that $hH = H$ because H is closed under its group operation. Then $g_2H = g_1hH = g_1H$. □

The above proposition says that the left cosets of H in G partition G into disjoint sets gH for $g \in G$. We can define the following relation on G:

$$g_1 \quad g_2 \quad ( ) \quad g_1H = g_2H \tag{3.1}$$

Using Proposition 3.15 it can be verified that this is an equivalence relation. We can then define the equivalence class of $g \in G$ as

$$[g] \stackrel{\text{def}}{=} \{g' \in G \mid g' \quad g\}.$$

It follows from (3.1) that

$$[g] = gH.$$

The set of all such equivalence classes is denoted as G/H, i.e.

$$G/H \stackrel{\text{def}}{=} \{gH \mid g \in G\}.$$

Definition 3.16. A subgroup N of a group G is called a normal subgroup of G if it is invariant under conjugation in G, that is,

$$\forall g \in G : gNg^{-1} = N.$$

Example 3.17. Let $G = \mathbb{Z}$ and $N = n\mathbb{Z}$ for $n \in \mathbb{Z}, n > 1$. Take any $g \in \mathbb{Z}$. Then we have

$$g + n\mathbb{Z} + (\quad g) = g + (\quad g) + n\mathbb{Z} = n\mathbb{Z},$$

which shows that $n\mathbb{Z}$ is a normal subgroup of $\mathbb{Z}$. Actually here, $G = \mathbb{Z}$ is a commutative group ($a + b = b + a$, $\forall a, b \in \mathbb{Z}$). That's why we can write

$$gNg^{-1} = gg^{-1}N = 1_GN = N$$

for $N = n\mathbb{Z}$ and $g \in \mathbb{Z}$. Commutativity of elements in $\mathbb{Z}$ allows us to change Ng^{-1} to $g^{-1}N$. That means, in general, that every subgroup of a commutative group is normal.

It turns out that if we define a normal subgroup N in this way, we then are able to turn the set of left cosets G/N into a group. It is crucial that the group law of G/N is induced from the group law of G.

Proposition 3.18. *Let G be a group and N be a normal subgroup of G. Define an operation on the set of left cosets, G/N, as follows:*

$$(aN)(bN) \stackrel{\text{def}}{=} (ab)N \tag{3.2}$$

Then:

1. *The operation (3.2) is well-defined.*
2. *The operation (3.2) turns G/N into a group.*

Proof. 1. To check that (3.2) is well-defined we need to show that the result doesn't depend on the choice of class representatives. Let $a^0N = aN$ and $b^0N = bN$. Then

$$(a^0N)(b^0N) = (a^0b^0)N = a^0(b^0N) = a^0(bN) = a^0(Nb) = (a^0N)b =$$

$$= (aN)b = a(Nb) = a(bN) = (ab)N = (aN)(bN).$$

2. The closure of G/N under this operation follows immediately from (3.2). This operation is associative, since

$$((aN)(bN))(cN) = ((ab)N)(cN) = ((ab)c)N = (a(bc))N =$$

$$= (aN)((bc)N) = (aN)((bN)(cN)).$$

The identity element is 1_GN and the inverse of aN is $a^{-1}N$.

□

For simplicity, we will write abN instead of $(ab)N$. According to Proposition 3.18 the following definition makes sence.

Definition 3.19. Let G be a group and N be a normal subgroup of G. We call G/N the quotient group of G by N.

The following proposition is known as the First Isomorphism Theorem for groups.

Proposition 3.20. *Let G and H be groups and let $\varphi\colon G \to H$ be a group homomorphism. Then:*

1. $\ker(\varphi)$ *is a normal subgroup of* G.
2. $\mathrm{im}(\varphi)$ *is a subgroup of* H.
3. *The quotient group* $G/\ker(\varphi)$ *is isomorphic to* $\mathrm{im}(\varphi)$.

In particular, if φ is surjective then H is isomorphic to $G/\ker(\varphi)$.

Proof. 1. Take $g \in G$. We need to prove that

$$gng^{-1} \in \ker(\varphi) \quad \forall n \in \ker(\varphi).$$

Applying φ to it we get

$$\varphi(gng^{-1}) = \varphi(g)\varphi(n)\varphi(g^{-1}) = \varphi(g)1_H\varphi(g^{-1}) = \varphi(g)\varphi(g^{-1}) = \varphi(gg^{-1}) = \varphi(1_G) = 1_H.$$

That means gng^{-1} lies in the kernel $\ker(\varphi)$.

2. We need to check that $\mathrm{im}(\varphi)$ is closed under multiplication in H. Take two elements $\varphi(g_1)$ and $\varphi(g_2)$ in the image $\mathrm{im}(\varphi)$. Then

$$\varphi(g_1)\varphi(g_2) = \varphi(g_1g_2) \in \mathrm{im}(\varphi).$$

3. Let $K = \ker(\varphi)$. Define a map

$$\begin{aligned} \overline{\varphi}\colon G/K &\to \operatorname{im}(\varphi) \\ gK &\mapsto \varphi(g) \end{aligned}$$

It is well-defined because if $g_1K = g_2K$ (which means $g_2^{-1}g_1 \in K$) then

$$\overline{\varphi}(g_1K) = \varphi(g_1) = \varphi(g_2g_2^{-1}g_1) = \varphi(g_2)\varphi(g_2^{-1}g_1) = \varphi(g_2)1_H = \varphi(g_2) = \overline{\varphi}(g_2K).$$

It is a group homomorphism since

$$\overline{\varphi}\big((g_1K)(g_2K)\big) = \overline{\varphi}(g_1g_2K) = \varphi(g_1g_2) = \varphi(g_1)\varphi(g_2) = \overline{\varphi}(g_1K)\overline{\varphi}(g_2K).$$

We show that $\overline{\varphi}$ is injective:

$$\overline{\varphi}(g_1K) = \overline{\varphi}(g_2K) \implies \varphi(g_1) = \varphi(g_2) \implies 1_H = (\varphi(g_1))^{-1}\varphi(g_2).$$

By Proposition 3.8,

$$1_H = (\varphi(g_1))^{-1}\varphi(g_2) = \varphi(g_1^{-1}g_2) \implies g_1^{-1}g_2 \in K \implies g_1K = g_2K.$$

Finally, we show that $\overline{\varphi}$ is surjective. But this is trivial because for any $\varphi(g) \in \operatorname{im}(\varphi)$ there is a coset gK which maps to $\varphi(g)$ by $\overline{\varphi}$. So $\overline{\varphi}$ is indeed an isomorphism. □

Remark 3.21. From now on, if we have a surjective group homomorphism $\varphi\colon G \to H$ with kernel $K \trianglelefteq G$, the induced isomorphism between the quotient group G/K and H will be denoted as $\overline{\varphi}\colon G/K \to H$ and we will write

$$G/K \overset{\overline{\varphi}}{=} H.$$

If a group homomorphism $\varphi\colon G \to H$ is injective, we write $\varphi\colon G \hookrightarrow H$.

Example 3.22. Let $G = \mathbb{Z}$ and $N = n\mathbb{Z}$ for some $n \in \mathbb{Z}, n > 1$. Let $H = \mathbb{Z}_n$. We can define a group homomorphism

$$\begin{aligned} \varphi\colon \mathbb{Z} &\to \mathbb{Z}_n \\ a &\mapsto a \bmod n \end{aligned}$$

We saw in Example 3.12 that $\ker(\varphi) = n\mathbb{Z}$ and $\operatorname{im}(\varphi) = \mathbb{Z}_n$. By Proposition 3.18, $n\mathbb{Z}$ is a normal subgroup of $\mathbb{Z}$ and

$$\mathbb{Z}/n\mathbb{Z} \overset{\overline{\varphi}}{=} \mathbb{Z}_n.$$

Proposition 3.23. *Let G be a group and N be a normal subgroup of G. If H is a subgroup of G such that $N \trianglelefteq H$, then N is a normal subgroup of H.*

Proof. Since N is normal in G, then

$$gNg^{-1} = N, \quad \forall g \in G.$$

In the above equation we can just take those g which are in H and obtain

$$gNg^{-1} = N, \quad \forall g \in H,$$

which means N is normal in H. □

In Chapter 5 we will work directly with objects which are isomorphic to some quotient group. That's why we are also interested in subgroups of quotient groups. We prove the following proposition which will be useful in Chapter 5.

Proposition 3.24. *Let G be a group and N, H be normal subgroups of G with $N \quad H$. Let G/H, G/N, H/N be the quotient groups. Then*

1. *H/N is a normal subgroup of G/N.*

2. *$(G/N)/(H/N) \overset{\overline{\beta}}{=} G/H$.*

Proof. We define a map:

$$\begin{aligned} \beta\colon G/N &\ ! \ G/H \\ gN &\ 7! \ gH \end{aligned}$$

We check that β is well-defined. Suppose $g_1N = g_2N$. Then $g_2{}^{1}g_1 \ 2 \ N$. Because $N \quad H$, then $g_2{}^{1}g_1 \ 2 \ H$, which means $g_1H = g_2H$. So, β is well-defined.
Now, β is a group homomorphism since

$$\beta\big((g_1N)(g_2N)\big) = \beta(g_1g_2N) = g_1g_2H = (g_1H)(g_2H) = \beta(g_1)\beta(g_2).$$

Obviously, β is surjective because for any coset gH just take a coset gN which maps to gH by β.
The kernel of β is:

$$\ker(\beta) = \{gN \ 2 \ G/N \mid \beta(gN) = 1_{G/H}\} = \{gN \ 2 \ G/N \mid gH = 1_G H\} =$$

$$= \{gN \ 2 \ G/N \mid g \ 2 \ H\} = H/N.$$

By Proposition 3.20, H/N is a normal subgroup of G/H and

$$(G/N)/(H/N) \overset{\overline{\beta}}{=} G/H \tag{3.3}$$

□

3.2 Permutation groups

Let X be a finite set. A bijective map from X to itself is called a permutation of X. Denote the set of all permutations of X as $\mathrm{S}(X)$. Then it is trivial to verify that $(\mathrm{S}(X), \ )$ forms a group where $\ $ is the operation of function composition. This group is finite of order $d!$, where d is the cardinality of X, since there are $d!$ different permutations of d elements. Sometimes, for simplification, we will write S_d instead of $\mathrm{S}(X)$. A subgroup $G \quad \mathrm{S}(X)$ is called a permutation group.

Remark 3.25. For $\sigma, \tau \ 2 \ \mathrm{S}(X)$, we will omit the sign $\ $ in their composition $\sigma \ \tau$ and write just $\sigma\tau$. In general, we will follows this rule for any two maps.

Example 3.26. Let $X = f1, 2, 3g$. Then $S(X)$ is the group of all permutations of X. There are $3! = 6$ di□erent permutations of 3 elements, so $jS(X)j = 6$. There is a subgroup $G \quad S(X)$ which consists of the identity permutation and the permutation which exchanges 1 and 2 and keeps 3 fixed. The order of G equals 2.

We will represent a permutation $\sigma \ 2 \ S(X)$ using cycle notation: for $X = f1, \ldots, dg$, the elements in each cycle are put inside parentheses, ordered so that $\sigma(j)$ immediately follows j or, if j is the last listed element of the cycle, then $\sigma(j)$ is the first element of the cycle. For example, if $X = f1, \ldots, 6g$ then the permutation

$$\sigma(1) = 4, \sigma(2) = 1, \sigma(3) = 3, \sigma(4) = 2, \sigma(5) = 6, \sigma(6) = 5$$

is

$$\sigma = (1\ 4\ 2)(3)(5\ 6)$$

in cycle notation.

Definition 3.27. A permutation group $G \quad S(X)$ is transitive if for every $x, x^0 \ 2 \ X$ there is an element $g \ 2 \ G$ such that $g(x) = x^0$.

Example 3.28. Let $G = h(1\ 3)(2\ 4), (1\ 4)(2\ 3)i \quad S_4$ be a subgroup. It is easy to see that it is transitive. If we let $G = h(1\ 2), (3\ 4)i \quad S_4$, then G is no longer transitive since there is no element $\sigma \ 2 \ G$ such that $\sigma(1) = 3$.

Definition 3.29. Let X be a finite set and $G \quad S(X)$ be a permutation group. For $x \ 2 \ X$, the stabilizer of x by G is the set

$$\mathrm{Stab}_G(x) = fg \ 2 \ G \ j \ g(x) = xg.$$

Proposition 3.30. *The stabilizer* $\mathrm{Stab}_G(x)$ *is a subgroup of* G.

Proof. Take $g_1, g_2 \ 2 \ \mathrm{Stab}_G(x)$. Then $g_1g_2(x) = g_1(g_2(x)) = g_1(x) = x$ and hence $g_1g_2 \ 2 \ \mathrm{Stab}_G(x)$. Obviously, $e_G \ 2 \ \mathrm{Stab}_G(x)$. If $g \ 2 \ \mathrm{Stab}_G(x)$ then $x = e_G(x) = g^{\ 1}g(x) = g^{\ 1}(g(x)) = g^{\ 1}(x)$, so $\mathrm{Stab}_G(x)$ is closed under taking inverses. □

Example 3.31. Let $X = f1, 2, 3, 4g$ and $G = h(1\ 3)(2\ 4), (1\ 4)(2\ 3)i \quad S_4$. Then

$$\mathrm{Stab}_G(1) = \mathrm{Stab}_G(2) = \mathrm{Stab}_G(3) = \mathrm{Stab}_G(4) = fe_Gg.$$

Example 3.32. Let $X = f1, 2, 3g$ and $G = S_3$. Then $\mathrm{Stab}_G(1) = fe_G, (2\ 3)g = h(2\ 3)i$.

Proposition 3.33. *Let* X *be a finite set and* $G \quad S(X)$ *a transitive permutation group. Fix an element* $x \ 2 \ X$ *and let* G_x *denotes the stabilizer* $\mathrm{Stab}_G(x)$. *Then the map*

$$\begin{aligned} \eta\colon G/G_x &\ ! \ X \\ gG_x &\ 7! \ g(x) \end{aligned}$$

is bijective.

Proof. We first show that η is well-defined. It means that if $g_1G_x = g_2G_x$, then $g_1(x) = g_2(x)$. The first equality means that $g_1 = g_2h$ for some h 2 G_x. Then $g_1(x) = g_2h(x) = g_2(h(x)) = g_2(x)$.

We that η is bijective. For injectivity, let $\eta(g_1G_x) = \eta(g_2G_x)$. Then $g_1(x) = g_2(x)$, or $x = g_1{}^{1}g_2(x)$. Hence $g_1{}^{1}g_2$ 2 G_x, or $g_2 = g_1h$ for some h 2 G_x. Then $g_1G_x = g_2G_x$. For surjectivity, by assumption G acts transitively on X, which means that for every x^{0} 2 X there is g 2 G with $x^{0} = g(x)$. Then $\eta(gG_x) = x^{0}$. □

3.3 Products of groups

Here we explain two basic definitions related to the notion of product of groups. We will refer to this section in Chapter 8.

Given two groups G and H, there are several ways how to construct a new group from G and H. The first one is analogous to the Cartesian product of two sets.

Definition 3.34. Given two groups $G = (G, \ )$ and $H = (H, \ )$, we define the direct product $G \quad H$ of G and H as follows:

1. The underlying set is the Cartesian product $G \quad H$, i.e.:

$$G \quad H = f(g,h) \ j \ g \ 2 \ G, h \ 2 \ Hg.$$

2. The operation $\star$ of $G \quad H$ is defined component-wise:

$$(g_1,h_1) \star (g_2,h_2) \overset{\text{def}}{=} (g_1 \quad g_2, h_1 \quad h_2)$$

Proposition 3.35. *The direct product $G \quad H$ constructed in Definition 3.34 satisfies the group axioms.*

Proof. It is obviously closed under the operation $\star$. The associativity of $\star$ follows from the associativity of and since:

$$((g_1,h_1) \star (g_2,h_2)) \star (g_3,h_3) = (g_1 \quad g_2, h_1 \quad h_2) \star (g_3,h_3) = ((g_1 \quad g_2) \quad g_3, (h_1 \quad h_2) \quad h_3) =$$

$$= (g_1 \quad (g_2 \quad g_3), h_1 \quad (h_2 \quad h_3)) = (g_1,h_1) \star (g_2 \quad g_3, h_2 \quad h_3) = (g_1,h_1) \star ((g_2,h_2) \star (g_3,h_3)).$$

The identity element is $(1_G, 1_H)$ since for every (g,h) 2 $G \quad H$:

$$(1_G,1_H) \star (g,h) = (1_G \quad g, 1_H \quad h) = (g,h) = (g \quad 1_G, h \quad 1_H) = (g,h) \star (1_G,1_H).$$

The inverse of (g,h) is $(g \ ^{1}, h \ ^{1})$ since

$$(g,h) \star (g \ ^{1}, h \ ^{1}) = (g \quad g \ ^{1}, h \quad h \ ^{1}) = (1_G,1_H) = (g \ ^{1} \quad g, h \ ^{1} \quad h) = (g \ ^{1}, h \ ^{1}) \star (g,h).$$

□

Example 3.36. Let $(\mathbb{R}, +)$ be the additive group of real numbers. Then the direct product of $(\mathbb{R}, +)$ with itself is the group $(\mathbb{R} \quad \mathbb{R}, \quad)$ with the operation given by

$$(a_1, b_1) \quad (a_2, b_2) = (a_1 + a_2, b_1 + b_2).$$

Let $(\mathbb{R}\ ,\ )$ be the multiplicative group of nonzero real numbers. The direct product of $(\mathbb{R}\ ,\ )$ with itself is the group $(\mathbb{R} \quad \mathbb{R}\ ,\ \quad)$ with the operation given by

$$(a_1, b_1) \quad (a_2, b_2) = (a_1 \quad a_2, b_1 \quad b_2).$$

We can also let $(\mathbb{R} \quad \mathbb{R}\ ,\ \quad)$ to be the direct product of $(\mathbb{R}, +)$ and $(\mathbb{R}\ ,\ )$. Then the operation is given by

$$(a_1, b_1) \quad (a_2, b_2) = (a_1 + a_2, b_1 \quad b_2).$$

As it was done in Section 3.1, we will write just the underlying set $G \quad H$ for the direct product of G and H.

There is also another way to construct a new group from the given two groups, which generalizes the direct product. However, first we need to be familiar with the notion of automorphism group of a group G.

Definition 3.37. Let G be a group. We define a set

$$\mathrm{Aut}(G) \stackrel{\text{def}}{=} \{\varphi \colon G \ ! \ G \mid \varphi \text{ is an isomorphism}\}.$$

It is easy to verify that such a set forms a group with respect to the operation of function composition . We call $(\mathrm{Aut}(G),\ )$ the automorphism group of G. The elements of $\mathrm{Aut}(G)$ are called the automorphisms of G.

Definition 3.38. Given two groups $G = (G,\ )$ and $H = (H,\ )$ and a group homomorphism $\varphi \colon H \ ! \ \mathrm{Aut}(G)$, we construct a new group $G \circ_\varphi H$, called the (outer) semidirect product of G and H with respect to φ, defined as follows:

1. The underlying set is the Cartesian product $G \quad H$.
2. The operation $\star$ is defined as:

$$(g_1, h_1) \star (g_2, h_2) \stackrel{\text{def}}{=} (g_1 \quad \varphi(h_1)(g_2), h_1 \quad h_2).$$

Proposition 3.39. *The semidirect product $G \circ_\varphi H$ constructed in Definition 3.38 satisfies the group axioms.*

Proof. It is obviously closed under $\star$ since $\varphi_{h_1} \stackrel{\text{def}}{=} \varphi(h_1)$ maps g_2 to some element in G. The associativity of $\star$ follows from the associativity of and since

$$((g_1, h_1) \star (g_2, h_2)) \star (g_3, h_3) = (g_1 \quad \varphi_{h_1}(g_2), h_1 \quad h_2) \star (g_3, h_3) =$$

$$= \Big(\big(g_1 \quad \varphi_{h_1}(g_2)\big) \quad \varphi_{h_1 \, h_2}(g_3), (h_1 \quad h_2) \quad h_3 \Big) =$$

$$= \Big(g_1 \quad \big(\varphi_{h_1}(g_2) \quad \varphi_{h_1 \, h_2}(g_3)\big), h1 \quad (h_2 \quad h_3) \Big) \stackrel{(1)}{=}$$

$$\overset{(1)}{=} \Big(g_1 \quad \big(\varphi_{h_1}(g_2) \quad \varphi_{h_1}(\varphi_{h_2}(g_3))\big), h1 \quad (h_2 \quad h_3)\Big) \overset{(2)}{=}$$
$$\overset{(2)}{=} \Big(g_1 \quad \varphi_{h_1}\big(g_2 \quad \varphi_{h_2}(g_3)\big), h1 \quad (h_2 \quad h_3)\Big) =$$
$$= (g_1, h_1) \star (g_2 \quad \varphi_{h_2}(g_3), h_2 \quad h_3) =$$
$$= (g_1, h_1) \star ((g_2, h_2) \star (g_3, h_3)),$$

where (1) and (2) follow from the facts that φ and φ_{h_1} are group homomorphisms. The identity element is $(1_G, 1_H)$ since for every (g, h) 2 G o$_\varphi$ H:

$$(1_G, 1_H) \star (g, h) = (1_G \quad \varphi_{1_H}(g), 1_H \quad h) \overset{(1)}{=} (g, h) \overset{(2)}{=} (g \quad \varphi_h(1_G), h \quad 1_H) = (g, h) \star (1_G, 1_H),$$

where both (1) and (2) follow from Proposition 3.8. The inverse of (g, h) is $(\varphi_{h\ 1}(g\ ^{1}), h\ ^{1})$ since

$$(g, h) \star (\varphi_{h\ 1}(g\ ^{1}), h\ ^{1}) = (g \quad \varphi_h(\varphi_{h\ 1}(g\ ^{1})), h \quad h\ ^{1}) = (g \quad \varphi_{h\,h\ 1}(g\ ^{1}), h \quad h\ ^{1}) =$$
$$= (g \quad \varphi_{1_H}(g\ ^{1}), 1_H) = (g \quad g\ ^{1}, 1_H) = (1_G, 1_H) = (\varphi_{h\ 1}(1_G), 1_H) =$$
$$= (\varphi_{h\ 1}(g\ ^{1} \quad g), h\ ^{1} \quad h) = (\varphi_{h\ 1}(g\ ^{1}) \quad \varphi_{h\ 1}(g), h\ ^{1} \quad h) = (\varphi_{h\ 1}(g\ ^{1}), h\ ^{1}) \star (g, h).$$

□

We note that the semidirect product is the generalization of the direct product (just take φ which sends all h 2 H to id_G). For simplicity, the semidirect product of groups G and H with respect to φ will be denoted as G o$_\varphi$ H.

Example 3.40. Let $K = \mathrm{S}(f a, b g)$ be the permutation group. Let G be the direct product $K \quad K \quad K$ and $H = \mathrm{S}(f 1, 2, 3 g)$. We define a map

$$\varphi \colon H \ ! \ \mathrm{Aut}(G)$$
$$h \ 7! \ \big(g \ 7! \ (g_{h\ 1(1)}, g_{h\ 1(2)}, g_{h\ 1(3)})\big)$$

where g_i denotes the i-th element of g. So, $\varphi_h = \varphi(h)$ just permutes the elements of g according to $h\ ^{1}$. We at first show that φ_h is a group homomorphism. Take g, g^{0} 2 G. Then

$$\varphi_h(g g^{0}) = \varphi_h\big(\underbrace{(g_1 g_1^{0}, g_2 g_2^{0}, g_3 g_3^{0})}_{\widetilde{g}}\big) = \varphi_h(\widetilde{g}) = (\widetilde{g}_{h\ 1(1)}, \widetilde{g}_{h\ 1(2)}, \widetilde{g}_{h\ 1(3)}) =$$
$$= (g_{h\ 1(1)} g_h^{0}{}_{\ 1(1)}, g_{h\ 1(2)} g_h^{0}{}_{\ 1(2)}, g_{h\ 1(3)} g_h^{0}{}_{\ 1(3)}) =$$
$$= (g_{h\ 1(1)}, g_{h\ 1(2)}, g_{h\ 1(3)})(g_h^{0}{}_{\ 1(1)}, g_h^{0}{}_{\ 1(2)}, g_h^{0}{}_{\ 1(3)}) = \varphi_h(g)\varphi_h(g^{0}).$$

Also φ_h is bijective since h is bijective. Therefore φ_h is an isomorphism. Now we show that φ is a group homomorphism. Take h_1, h_2 2 H and let $\widetilde{g} = \varphi_{h_2}(g) = (g_{h_2\ 1(1)}, g_{h_2\ 1(2)}, g_{h_2\ 1(3)})$. Then

$$\varphi(h_1 h_2) = \big(g \ 7! \ (g_{(h_1 h_2)\ 1(1)}, g_{(h_1 h_2)\ 1(2)}, g_{(h_1 h_2)\ 1(3)})\big) =$$
$$= \big(g \ 7! \ (g_{h_2\ 1(h_1\ 1(1))}, g_{h_2\ 1(h_1\ 1(2))}, g_{h_2\ 1(h_1\ 1(3))})\big) \overset{(1)}{=} \big(g \ 7! \ (\widetilde{g}_{h_1\ 1(1)}, \widetilde{g}_{h_1\ 1(2)}, \widetilde{g}_{h_1\ 1(3)})\big) =$$

$$= \big(g \circ (\varphi_{h_1}(\widetilde{g})\big) = \big(g \circ (\varphi_{h_1}(\varphi_{h_2}(g))\big) = \varphi(h_1)\varphi(h_2).$$

Equality (1) is true because $\widetilde{g}_i = g_{h_2^{-1}(i)}$ for all $i = 1, 2, 3$. Thus, we may construct the semidirect product $G \wr_\varphi H$. At first, the order of $G \wr_\varphi H$ equals $2!^3 \cdot 3! = 48$. To understand the structure of $G \wr_\varphi H$ we will embed it into $S(\underbrace{f(a,1),(b,1),(a,2),(b,2),(a,3),(b,3)g}_{X})$.

Define a map

$$\begin{aligned} \psi\colon G \wr_\varphi H &\to S(X) \\ (g,h) &\mapsto \big((i,j) \mapsto (g_{h(j)}(i), h(j))\big) \end{aligned}$$

The map $\psi\big((g,h)\big)$ is bijective since $g_i, i = 1, 2, 3$ and h are bijective. We prove that ψ is a group homomorphism. Take (g,h) and (g^0,h^0) in $G \wr_\varphi H$ and let $\widetilde{g} = \varphi_h(g^0)$. Then

$$\psi\big((g,h)(g^0,h^0)\big) = \psi\big((g\widetilde{g}, hh^0)\big) =$$

$$= \big((i,j) \mapsto ((g\widetilde{g})_{(hh^0)(j)}(i), (hh^0)(j))\big) =$$

$$= \big((i,j) \mapsto ((g\widetilde{g})_{h(h^0(j))}(i), h(h^0(j)))\big) =$$

$$= \big((i,j) \mapsto ((g_{h(h^0(j))}\widetilde{g}_{h(h^0(j))})(i), h(h^0(j)))\big) =$$

$$= \big((i,j) \mapsto ((g_{h(h^0(j))}g^0_{h^{-1}(h(h^0(j)))})(i), h(h^0(j)))\big) =$$

$$= \big((i,j) \mapsto ((g_{h(h^0(j))}g^0_{h^0(j)})(i), h(h^0(j)))\big) =$$

$$= \big((i,j) \mapsto (g_{h(h^0(j))}(\underbrace{g^0_{h^0(j)}(i)}_{i^0}), h(\underbrace{h^0(j)}_{j^0}))\big) =$$

$$= \big((i,j) \mapsto (g_{h(j^0)}(i^0), h(j^0))\big) =$$

$$= \big((i,j) \mapsto \psi_{(g,h)}((i^0,j^0))\big) =$$

$$= \big((i,j) \mapsto \psi_{(g,h)}(\psi_{(g^0,h^0)}((i,j)))\big) =$$

$$= \psi\big((g,h)\big)\psi\big((g^0,h^0)\big).$$

The group homomorphism ψ is injective since h and $g_i, i = 1, 2, 3$ are injective. To show how the elements of $\mathrm{im}(\psi)$ act on X let's take, for example, $g = \big((a\ b), 1_K, 1_K\big) \in G$ and $h = (1\ 2\ 3) \in H$, where $(a\ b) \in K$ permutes a and b. Then the action of $\psi\big((g,h)\big)$ on X can be divided into 2 stages:

$$\begin{bmatrix}(a,1)\\(b,1)\\(a,2)\\(b,2)\\(a,3)\\(b,3)\end{bmatrix} \mapsto \begin{bmatrix}(b,1)\\(a,1)\\(a,2)\\(b,2)\\(a,3)\\(b,3)\end{bmatrix} \mapsto \begin{bmatrix}(a,2)\\(b,2)\\(a,3)\\(b,3)\\(b,1)\\(a,1)\end{bmatrix}.$$

Notice that every element $f \in \mathrm{im}(\psi)$ respects the partition of X into blocks:

$$B = \Big\{\big\{(a,1),(b,1)\big\}, \big\{(a,2),(b,2)\big\}, \big\{(a,3),(b,3)\big\}\Big\}.$$

This means that

$$f(B) = B \text{ or } f(B) \setminus B = ? \quad 8B \in \mathcal{B},\ 8f \in \mathrm{im}(\psi).$$

If such a partition $\mathcal{B}$ of X exists then we say that $\mathrm{im}(\psi)$ acts imprimitively on X (see Chapter 8). The group $G \wr_\varphi H = (K \times \cdots \times K) \rtimes_\varphi H$ is denoted in the literature as $K \wr H$ (or K wr H) and is called the wreath product of K and H.

3.4 Relation between the stabilizer, normalizer and centralizer

In this section we would like to reveal the connection between the stabilizer, normalizer and centralizer inside the permutation group, since this will play a key role in Chapter 5. We already know from Section 3.2 what the stabilizer of an element is. We give the following two definitions.

Definition 3.41. Let G be a group. The normalizer of a subset $S \subseteq G$ is defined as

$$\mathrm{N}_G(S) \stackrel{\text{def}}{=} \{g \in G \mid gS = Sg\}.$$

We will be mostly interested in normalizers of subgroups.

Example 3.42. Let $X = \{1, 2, 3, 4\}$ and $G = \langle (1\ 2)(3\ 4), (1\ 3) \rangle \leq \mathrm{S}(X)$. Take $x = 1 \in X$. Using GAP we can compute the stabilizer of x in G. The command which does that is

```
Stab := Stabilizer(G,1);
```

We obtain

$$\mathrm{Stab}_G(x) = \langle (2\ 4) \rangle.$$

Similarly, the normalizer of $\mathrm{Stab}_G(x)$ in G can be computed using the command

```
Norm := Normalizer(G,Stab);
```

We obtain

$$\mathrm{N}_G(\mathrm{Stab}_G(x)) = \langle (2\ 4), (1\ 3) \rangle.$$

Notice that if H is a subgroup of G, then from Definition 3.41 it follows that H is a normal subgroup of $\mathrm{N}_G(H)$ and, thus, we may construct the quotient group $\mathrm{N}_G(H)/H$. It is obviously isomorphic to $\mathbb{Z}/2\mathbb{Z}$.

Definition 3.43. Let G be a group. The centralizer of a subset $S \subseteq G$ is defined as

$$\mathrm{C}_G(S) \stackrel{\text{def}}{=} \{g \in G \mid gs = sg \text{ for all } s \in S\}.$$

We will be mostly interested in centralizers of subgroups.

Example 3.44. Let $X = \{1, 2, 3, 4\}$ and $G = \langle (1\ 2)(3\ 4), (1\ 3) \rangle \leq \mathrm{S}(X)$ be a permutation group. Again, we use GAP to compute the centralizer of G in $\mathrm{S}(X)$:

```
Cent := Centralizer(G,SymmetricGroup(4));
```

We obtain

$$\mathrm{C}_{\mathsf{S}(X)}(G) = \langle (1\ 3)(2\ 4) \rangle.$$

Consider a finite set X and a transitive permutation group $G \leq \mathsf{S}(X)$. Take some $x \in X$ and denote $G_x = \mathrm{Stab}_G(x)$. Basically, in this section we prove that

$$\mathrm{N}_G(G_x)/G_x \overset{\varrho}{=} \mathrm{C}_{\mathsf{S}(X)}(G) \tag{3.4}$$

Example 3.45. Let $X = \{1,2,3,4\}$ and $G = \langle (1\ 2)(3\ 4), (1\ 3) \rangle \leq \mathsf{S}(X)$ be a permutation group. Take $x = 1 \in X$. Then, according to Examples 3.42 and 3.44, we have

$$\mathrm{N}_G(G_x)/G_x = \mathsf{Z}/2\mathsf{Z} = \mathrm{C}_{\mathsf{S}(X)}(G).$$

We now prove relation (3.4). Let G/G_x be the set of left cosets. Then we can define a group homomorphism

$$\begin{aligned} \rho_L \colon G &\to \mathsf{S}(G/G_x) \\ g' &\mapsto \left(gG_x \mapsto g'gG_x\right) \end{aligned}$$

It is easy to see that for every $g' \in G$ the map $\rho_L(g')$ is well-defined since if $g_1G_x = g_2G_x$, then $g'g_1G_x = g'g_2G_x$. So, ρ_L is well-defined. It is a group homomorphism since

$$\rho_L(g'g'') = \left(gG_x \mapsto (g'g'')gG_x\right) = \left(gG_x \mapsto g'(g''g)G_x\right) = \rho_L(g')\rho_L(g'').$$

We can try to define another group homomorphism

$$\begin{aligned} \rho_R \colon G &\to \mathsf{S}(G/G_x) \\ n &\mapsto \left(gG_x \mapsto gn^{-1}G_x\right) \end{aligned}$$

However, we will not succeed here because the map $\rho_R(n)$ will not be well-defined for all $n \in G$. For $\rho_R(n)$ to be well-defined means

$$g_1G_x = g_2G_x \Rightarrow g_1n^{-1}G_x = g_2n^{-1}G_x, \quad \forall g_1, g_2 \in G. \tag{3.5}$$

We can see that for $g_1, g_2 \in G$,

$$g_1G_x = g_2G_x \iff g_2^{-1}g_1G_x = G_x \iff g_2^{-1}g_1 \in G_x,$$
$$g_1n^{-1}G_x = g_2n^{-1}G_x \iff ng_2^{-1}g_1n^{-1}G_x = G_x \iff ng_2^{-1}g_1n^{-1} \in G_x.$$

Thus, (3.5) can be rewritten as

$$g_2^{-1}g_1 \in G_x \Rightarrow ng_2^{-1}g_1n^{-1} \in G_x, \quad \forall g_1, g_2 \in G. \tag{3.6}$$

As we vary g_1 in G_x and keep $g_2 = 1_G$ fixed, the values $g_2^{-1}g_1 = g_1$ run through all the elements of G_x. Thus, by Definition 3.41, we can equivalently rewrite (3.6) as $n \in \mathrm{N}_G(G_x)$. Thus, $\rho_R(n)$ is well-defined if and only if $n \in NG_x \overset{\text{def}}{=} \mathrm{N}_G(G_x)$. So, we redefine ρ_R to be

$$\begin{aligned} \rho_R \colon NG_x &\to \mathsf{S}(G/G_x) \\ n &\mapsto \left(gG_x \mapsto gn^{-1}G_x\right) \end{aligned}$$

Also notice that ρ_R is a group homomorphism because

$$\rho_R(n_1n_2) = \Big(gG_x \mapsto g(n_1n_2)\ {}^1G_x\Big) = \Big(gG_x \mapsto g(n_2\,{}^1n_1\,{}^1)G_x\Big) =$$

$$= \Big(gG_x \mapsto (gn_2\,{}^1)n_1\,{}^1G_x\Big) = \rho_R(n_1)\rho_R(n_2).$$

We can create the following diagram (ϕ is to be defined):

$$\begin{array}{ccccc} G & \xrightarrow{\rho_L} & \mathrm{S}(G/G_x) & \xleftarrow{\rho_R} & NG_x \\ & \searrow^{\mathrm{id}_G} & \downarrow \phi & & \\ & & \mathrm{S}(X) & & \end{array} \tag{3.7}$$

Recall the map $\eta\colon G/G_x \to X$ from Proposition 3.33. We define ϕ to be

$$\begin{aligned} \phi\colon \mathrm{S}(G/G_x) &\to \mathrm{S}(X) \\ \sigma &\mapsto \Big(\tau\colon x^0 \mapsto \eta(\sigma(\eta\ {}^1(x^0)))\Big) \end{aligned}$$

(What τ does is just pulling back x^0 to the coset gG_x with $x^0 = g(x)$, apply σ to this coset and map it back to X by η.) It is true that ϕ is a group homomorphism because

$$\phi(\sigma_1\sigma_2) = \Big(x^0 \mapsto \eta(\sigma_1\sigma_2(\eta\ {}^1(x^0)))\Big) =$$

$$= \Big(x^0 \mapsto \eta(\sigma_1(\eta\ {}^1(\eta(\sigma_2(\eta\ {}^1(x^0))))))\Big)$$

$$= \Big(x^0 \mapsto \phi(\sigma_1)\big(\phi(\sigma_2)(x^0)\big)\Big) = \phi(\sigma_1)\phi(\sigma_2).$$

Also ϕ is bijective since η is bijective. Thus, ϕ is an isomorphism.

Proposition 3.46. *Diagram* (3.7) *is commutative which means*

$$\phi\rho_L = \mathrm{id}_G.$$

Proof. Take any $g^0 \in G$. Then $\tau = \phi(\rho_L(g^0)) \in \mathrm{S}(X)$ can be written as

$$\tau\colon x^0 \mapsto g^0g(x),$$

where $x^0 = g(x)$. So, it is the same as

$$\tau\colon x^0 \mapsto g^0(x^0),$$

since $g^0g(x) = g^0(g(x)) = g^0(x^0)$. But this is exactly what $\mathrm{id}_G(g^0) = g^0$ does. □

Let's talk now about injectivity of ρ_L and ρ_R. If ρ_L wasn't injective then $\phi\rho_L = \mathrm{id}_G$ wouldn't be injective, which is impossible. Thus, ρ_L is injective. So, diagram (3.7) now looks like

$$\begin{array}{ccccc} G & \xhookrightarrow{\rho_L} & \mathrm{S}(G/G_x) & \xleftarrow{\rho_R} & NG_x \\ & \searrow^{\mathrm{id}_G} & = \downarrow \phi & & \\ & & \mathrm{S}(X) & & \end{array}$$

Now we claim that $\ker(\rho_R) = G_x$. Indeed, take any $n \in G_x$. Then

$$\rho_R(n) = \Big(gG_x \mapsto gn\ {}^1G_x\Big) = \Big(gG_x \mapsto gG_x\Big) = 1_{\mathsf{S}(G/G_x)}.$$

Conversely, suppose $n \in \ker(\rho_R)$. Then $\rho_R(n)$ sends $1_G G_x$ to $1_G n\ {}^1G_x = 1_G G_x$. But this means $n \in G_x$. Hence, according to Proposition 3.20, the diagram can be redrawn as

$$\begin{array}{ccccc} G & \xrightarrow{\rho_L} & \mathsf{S}(G/G_x) & \xleftarrow{\overline{\rho_R}} & NG_x/G_x \\ & \searrow^{\mathsf{id}_G} & \Big\downarrow{\scriptstyle \phi} & & \\ & & \mathsf{S}(X) & & \end{array} \tag{3.8}$$

We now give the main proposition of this chapter.

Proposition 3.47. *In diagram* (3.8) *it holds that*

$$\mathrm{C}_{\mathsf{S}(G/G_x)}\left(\rho_L(G)\right) = \overline{\rho_R}(NG_x/G_x) \tag{3.9}$$

Proof. $\supseteq$: Take $\Big(\sigma_R\colon gG_x \mapsto gn\ {}^1G_x\Big) \in \overline{\rho_R}(NG_x/G_x)$ and $\Big(\sigma_L\colon gG_x \mapsto g'gG_x\Big) \in \rho_L(G)$. We need to prove that $\sigma_R\sigma_L = \sigma_L\sigma_R$, or that

$$\Big(\sigma_R\colon gG_x \mapsto gn\ {}^1G_x\Big)\Big(\sigma_L\colon gG_x \mapsto g'gG_x\Big) =$$

$$= \Big(\sigma_L\colon gG_x \mapsto g'gG_x\Big)\Big(\sigma_R\colon gG_x \mapsto gn\ {}^1G_x\Big),$$

which is equivalent to showing that

$$\Big(gG_x \mapsto (g'g)n\ {}^1G_x\Big) = \Big(gG_x \mapsto g'(gn\ {}^1)G_x\Big).$$

The latter equality obviously follows from the associativity law.

$\subseteq$: Take $\sigma \in \mathrm{C}_{\mathsf{S}(G/G_x)}\left(\rho_L(G)\right)$. Then

$$\sigma\sigma_L = \sigma_L\sigma, \quad \forall \sigma_L \in \rho_L(G).$$

It means that

$$\Big(gG_x \mapsto \sigma(g'gG_x)\Big) = \Big(gG_x \mapsto g'(\sigma(gG_x))\Big),$$

which is equivalent to

$$\sigma(g'gG_x) = g'(\sigma(gG_x)), \quad \forall g, g' \in G.$$

Take $gG_x = 1_G G_x$. Then

$$\sigma(g'G_x) = g'(\sigma(1_G G_x)), \quad \forall g' \in G.$$

If we denote $\sigma(1_G G_x) = n\ {}^1G_x$ for some $n \in G$, then we see that

$$\sigma(g'G_x) = g'n\ {}^1G_x, \quad \forall g' \in G.$$

It remains to prove that $n \in NG_x$. But it can be proven in exactly the same way as in (3.6) using the following fact (which says that σ is well-defined):

$$g'G_x = g''G_x \ \Rightarrow\ g'n\ {}^1G_x = g''n\ {}^1G_x.$$

So, σ indeed lies in $\overline{\rho_R}(NG_x/G_x)$. □

Relation (3.4) can be proved from (3.9) by showing that

$$\mathrm{C}_{\mathrm{S}(G/G_x)}\left(\rho_L(G)\right) = \mathrm{C}_{\mathrm{S}(X)}\left(G\right),$$

which is "kind of obvious" (for the proof see Proposition 5.26).

4 Elements of Algebraic Geometry

Every system of polynomial equations defines an object called a□ne variety. The branch of science which studies a□ne varieties is called algebraic geometry. In this chapter we are going to explain basic elements of algebraic geometry (i.e. polynomial ideals, a□ne varieties, the Zariski topology, etc.). The final concept which we will be interested in is the concept of rational maps between a□ne varieties. The material described here will be further used in Chapter 7.

4.1 Affine varieties

The set of all polynomials in $x_1, \ldots, x_n$ with coe□cients in C is denoted $\mathsf{C}[x_1, \ldots, x_n]$. One can show that $\mathsf{C}[x_1, \ldots, x_n]$ is a (commutative) ring.

Definition 4.1. Let $f_1, \ldots, f_m$ be polynomials in $\mathsf{C}[x_1, \ldots, x_n]$. Then we set

$$\mathbf{V}(f_1, \ldots, f_m) = f(a_1, \ldots, a_n) \; 2 \; \mathsf{C}^n \; j \; f_i(a_1, \ldots, a_n) = 0 \text{ for all } 1 \quad i \quad mg.$$

We call $\mathbf{V}(f_1, \ldots, f_m)$ the a□ne variety defined by $f_1, \ldots, f_m$.

In other words, $\mathbf{V}(f_1, \ldots, f_m)$ is the solution set of the polynomial system defined by $f_1, \ldots, f_m$.

Example 4.2. Let $f_1 = x \quad 1$ and $f_2 = y \quad 2$ be the polynomials in $\mathsf{C}[x, y]$. Then

$$\mathbf{V}(f_1, f_2) = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}.$$

Example 4.3. Let $f_1 = x^3 \quad 1$ and $f_2 = xy \quad 1$. From $f_1 = 0$ we get that x equals $1, e^{2\pi i \frac{1}{3}}, e^{2\pi i \frac{2}{3}}$. Substituting that into $f_2 = 0$ we get $y = \frac{1}{x}$ equals $1, e^{2\pi i \frac{2}{3}}, e^{2\pi i \frac{1}{3}}$. Hence

$$\mathbf{V}(f_1, f_2) = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} e^{2\pi i \frac{1}{3}} \\ e^{2\pi i \frac{2}{3}} \end{bmatrix}, \begin{bmatrix} e^{2\pi i \frac{2}{3}} \\ e^{2\pi i \frac{1}{3}} \end{bmatrix} \right\}.$$

Example 4.4. Let $f = x^2 + y^2 \quad 1 \; 2 \; \mathsf{C}[x, y]$. It is well-known that over the real numbers f defines a circle. Over the complex numbers there are more points in $\mathbf{V}(f)$. We can see that $\mathbf{V}(f)$ is infinite since C is infinite.

We already know from Definition 2.2 what an ideal is. Since $\mathbb{C}[x_1, \ldots, x_n]$ is a ring, we can study its ideals. We already know from Proposition 2.4 that for $f_1, \ldots, f_m$ the set of all polynomial combinations $\langle f_1, \ldots, f_m \rangle$ is an ideal of $\mathbb{C}[x_1, \ldots, x_n]$.

It turns out that to every affine variety $X \subseteq \mathbb{C}^n$ we can associate an ideal in the following way. We define

$$\mathbf{I}(X) = \{f \mid f \in \mathbb{C}[x_1, \ldots, x_n] \mid f(a_1, \ldots, a_n) = 0 \text{ for all } (a_1, \ldots, a_n) \in X\}.$$

If $X = \emptyset$, we let $\mathbf{I}(X) = \mathbb{C}[x_1, \ldots, x_n]$. The crucial observation is that $\mathbf{I}(X)$ is an ideal.

Proposition 4.5. *If $X \subseteq \mathbb{C}^n$ is an affine variety, then $\mathbf{I}(X) \subseteq \mathbb{C}[x_1, \ldots, x_n]$ is an ideal. We call $\mathbf{I}(X)$ the ideal of X.*

Proof. It is obvious that $0 \in \mathbf{I}(X)$ since the zero polynomial vanishes on all of $\mathbb{C}^n$, and so, in particular it vanishes on X. Next, suppose $f, g \in \mathbf{I}(X)$ and $h \in \mathbb{C}[x_1, \ldots, x_n]$. Let $(a_1, \ldots, a_n)$ be an arbitrary point of X. Then

$$f(a_1, \ldots, a_n) + g(a_1, \ldots, a_n) = 0 + 0 = 0,$$

$$h(a_1, \ldots, a_n) f(a_1, \ldots, a_n) = h(a_1, \ldots, a_n) \cdot 0 = 0,$$

and it follows that $\mathbf{I}(X)$ is an ideal. □

Proposition 4.6. *Every ideal $I \subseteq \mathbb{C}[x_1, \ldots, x_n]$ is finitely generated. In other words, $I = \langle f_1, \ldots, f_m \rangle$ for some $f_1, \ldots, f_m \in I$.*

Proof. [6, Chapter 2, §5, Theorem 4]. □

4.2 Regular and rational functions

Let $X \subseteq \mathbb{C}^n$ be an affine variety.

Definition 4.7. A function $f : X \to \mathbb{C}$ is said to be regular if there exists a polynomial F with coefficients in $\mathbb{C}$ such that $f(p) = F(p)$ for all $p \in X$.

The set of all regular functions on X is denoted by $\mathbb{C}[X]$. We can define a ring structure on the set $\mathbb{C}[X]$: the operations $f_1 + f_2$ and $f_1 \cdot f_2$ can be defined as

$$(f_1 + f_2)(p) \stackrel{\text{def}}{=} f_1(p) + f_2(p),$$

$$(f_1 \cdot f_2)(p) \stackrel{\text{def}}{=} f_1(p) \cdot f_2(p)$$

for every $p \in X$. It is then straightforward to check that such a definition turns $\mathbb{C}[X]$ into a ring, which we will call the coordinate ring of X.

It is easy to see that for a given function $f \in \mathbb{C}[X]$ its defining polynomial F is not unique: we can add to F any other polynomial G which vanishes on X without altering f. This can be explained algebraically as follows. Consider the map

$$\begin{aligned} \phi : \mathbb{C}[x_1, \ldots, x_n] &\to \mathbb{C}[X] \\ F &\mapsto (p \mapsto F(p)) \end{aligned} \tag{4.1}$$

It is obvious that ϕ is a ring homomorphism. It is surjective, because any regular function on X is defined by some polynomial. Its kernel $\ker(\phi)$ consists of polynomials which vanish on X. By definition, $\ker(\phi) = \mathbf{I}(X)$. Then, by Proposition 2.14, we have

$$\mathsf{C}[x_1, \ldots, x_n]/\mathbf{I}(X) \overset{\overline{\phi}}{=} \mathsf{C}[X] \tag{4.2}$$

So, every function on X can by identified with a coset $F + \mathbf{I}(X)$ for some polynomial F.

Example 4.8. Consider $X = f \; ig \quad \mathsf{C}$ given by the polynomial $f(x) = x^2 + 1$. Then $\mathbf{I}(X) = hx^2 + 1i \quad \mathsf{C}[x]$ and

$$\mathsf{C}[x]/hx^2 + 1i \overset{\overline{\phi}}{=} \mathsf{C}[X].$$

Because $x^2 + 1$ is reducible over C, the resulting coordinate ring has nontrivial zero divisors $[x + i] = (x + i) + hx^2 + 1i$ and $[x \quad i] = (x \quad i) + hx^2 + 1i$, and hence is not an integral domain.

Definition 4.9. An a□ne variety $X \quad \mathsf{C}^n$ is irreducible if whenever X is written in the form $X = X_1 \; [\; X_2$, where X_1 and X_2 are a□ne varieties, then either $X_1 = X$ or $X_2 = X$.

Example 4.10. Let $X = \mathbf{V}(xz, yz) \quad \mathsf{C}^3$. We can write $X = X_1 \; [\; X_2$ for $X_1 = \mathbf{V}(x, y)$ and $X_2 = \mathbf{V}(z)$. Thus, X is not irreducible. Let $Y = \mathbf{V}(y \quad x^2) \quad \mathsf{C}^2$. It turns out that Y is irreducible. However, it is hard to prove that directly from Definition 4.9. Proposition 4.11 turns this question into an algebraic problem.

Proposition 4.11. *Let $X \quad \mathsf{C}^n$ be an a ne variety. Then X is irreducible if and only if $\mathbf{I}(X)$ is a prime ideal.*

Proof. [6, Chapter 4, §5, Proposition 3]. □

Using isomorphism (4.2), Proposition 4.11 and Proposition 2.27 we obtain the following corollary.

Corollary 4.12. *For a nonempty irreducible a ne variety $X \quad \mathsf{C}^n$, its coordinate ring $\mathsf{C}[X]$ is an integral domain.*

Let $X \quad \mathsf{C}^n$ be an irreducible a□ne variety. Hence, using Corollary 4.12 and Proposition 2.22, the following definition makes sense.

Definition 4.13. The function field of a nonempty irreducible a□ne variety $X \quad \mathsf{C}^n$ is the field of fractions of $\mathsf{C}[X]$. This field is denoted as $\mathsf{C}(X)$.

So, by Proposition 2.22 and Remark 2.23, the elements of $\mathsf{C}(X)$ are the equivalence classes $\left[\frac{f}{g}\right]$, where $f, g \; 2 \; \mathsf{C}[X]$, $g \; 6= \; 0$.

Definition 4.14. Let $X \quad \mathsf{C}^n$ be an irreducible a□ne variety. A rational function on X is an element $f \; 2 \; \mathsf{C}(X)$.

Remark 4.15. Given an irreducible a□nem variety $X \quad \mathsf{C}^n$ and a rational function $\varphi = \left[\frac{f}{g}\right] \; 2 \; \mathsf{C}(X)$ we can try to define the evaluation of φ at points of X as $\varphi(p) = \frac{f(p)}{g(p)}$. However, we may not succeed always: there may exist some points $p \; 2 \; X$ for which $g(p) = 0$.

Definition 4.16. A rational function φ 2 C(X) is said to be regular at p 2 X if it has a representative $\frac{f}{g}$ with $g(p)$ 6= 0.

Example 4.17. Let $X = \mathbf{V}(xy \quad z^2) \quad$ C^3. The ideal $\mathbf{I}(X)$ is prime, since $xy \quad z^2$ is irreducible over C. So, by Proposition 4.11, X is irreducible. Hence the function field C(X) can be constructed. Take a rational function $\varphi = \left[\frac{x}{z}\right]$ 2 C(X). (For simplicity we write just $\left[\frac{x}{z}\right]$ instead of $\left[\frac{\phi(x)}{\phi(z)}\right]$ for ϕ from (4.1).) Me may notice that $\left[\frac{z}{y}\right]$ 2 C(X) defines the same rational function because

$$xy \quad z^2 = 0 \text{ 2 C}[X] \ ) \ \frac{x}{z} \quad \frac{z}{y}.$$

Take a point $p = (0, 1, 0)$ 2 X. If we try to evaluate $\frac{x}{z}$ at p we will not succeed, as $z(p) = 0$. But we can evaluate $\frac{z}{y}$ on p and obtain $\frac{z(p)}{y(p)} = \frac{0}{1} = 0$. By Definition 4.16, we claim that $\varphi = \left[\frac{x}{z}\right] = \left[\frac{z}{y}\right]$ is regular at $p = (0, 1, 0)$ and $\varphi(p) = 0$.

However, there is still one thing which hasn't been explained yet: why two di□erent representatives $\frac{f_1}{g_1}, \frac{f_2}{g_2}$ of the same rational function φ 2 C(X) should give the same value after evaluation at some p 2 X, for which $g_1(p), g_2(p)$ 6= 0? As $\frac{f_1}{g_1}$ and $\frac{f_2}{g_2}$ represent the same rational function φ, we should have $f_1 g_2 \quad f_2 g_1 = 0$ 2 C$[X]$. Then

$$f_1 g_2 \quad f_2 g_1 = 0 \ ) \ f_1(p) g_2(p) \quad f_2(p) g_1(p) = (f_1 g_2 \quad f_2 g_1)(p) = 0(p) = 0 \ ) \ \frac{f_1(p)}{g_1(p)} = \frac{f_2(p)}{g_2(p)}.$$

Hence, if a rational function φ 2 C(X) is regular at p 2 X, then its value at p doesn't depend on the choice of representative $\frac{f}{g}$ for which $g(p)$ 6= 0.

4.3 Subvarieties

Definition 4.18. Let $X \quad$ C^n be an a□ne variety.

(i) For $f_1, \ldots, f_m$ 2 C$[X]$ we define

$$\mathbf{V}_X(f_1, \ldots, f_m) = \{p \text{ 2 } X \mid f_i(p) = 0 \text{ for all } 1 \quad i \quad m\}.$$

We call $\mathbf{V}_X(f_1, \ldots, f_m)$ the subvariety of X.

(ii) For an ideal $J \quad$ C$[X]$ we define

$$\mathbf{V}_X(J) = \{p \text{ 2 } X \mid f(p) = 0 \text{ for all } f \text{ 2 } J\}.$$

(iii) For each subset $Y \quad X$, we define

$$\mathbf{I}_X(Y) = \{f \text{ 2 C}[X] \mid f(p) = 0 \text{ for all } p \text{ 2 } Y\}.$$

Proposition 4.19. *Let X be an a ne variety and Y be a subvariety of X. Then* $\mathbf{I}_X(Y)$ C$[X]$ *is an ideal.*

Proof. It is obvious that $0 \in \mathbf{I}_X(Y)$ since the zero function on X vanishes also on the subset $Y \subset X$. Next, suppose $f, g \in \mathbf{I}_X(Y)$ and $h \in \mathbb{C}[X]$. Take a point $p \in Y$. Then

$$f(p) + g(p) = 0 + 0 = 0,$$

$$h(p)f(p) = h(p) \cdot 0 = 0,$$

and it follows that $\mathbf{I}_X(Y)$ is an ideal. □

Proposition 4.20. *Let X be an affine variety and $f_1, \ldots, f_m \in \mathbb{C}[X]$. Then*

$$\mathbf{V}_X(f_1, \ldots, f_m) = \mathbf{V}_X(\langle f_1, \ldots, f_m \rangle).$$

Proof. $\subseteq$: Take $p \in \mathbf{V}_X(\langle f_1, \ldots, f_m \rangle)$. Since $f_1, \ldots, f_m \in \langle f_1, \ldots, f_m \rangle$, then $f_1(p) = \cdots = f_m(p) = 0$. Hence $p \in \mathbf{V}_X(f_1, \ldots, f_m)$.

$\supseteq$: Take $p \in \mathbf{V}_X(f_1, \ldots, f_m)$ and $f \in \langle f_1, \ldots, f_m \rangle$. Then $f = \sum_{i=1}^m g_i f_i$ for $g_1, \ldots, g_m \in \mathbb{C}[X]$. Evaluating f at p we obtain

$$f(p) = \sum_{i=1}^m g_i(p) f_i(p) = \sum_{i=1}^m g_i(p) \cdot 0 = 0.$$

Then $p \in \mathbf{V}_X(\langle f_1, \ldots, f_m \rangle)$. □

Let X be an affine variety defined by $F_1, \ldots, F_m \in \mathbb{C}[x_1, \ldots, x_n]$ and $Y = \mathbf{V}_X(g_1, \ldots, g_r)$ be a subvariety of X for $g_1, \ldots, g_r \in \mathbb{C}[X]$. It is easy to see that $Y \subset \mathbb{C}^n$ is an affine variety. Indeed, if we take representatives $G_1, \ldots, G_r \in \mathbb{C}[x_1, \ldots, x_n]$ of $g_1, \ldots, g_r$, then

$$Y = \mathbf{V}(F_1, \ldots, F_m, G_1, \ldots, G_r).$$

Example 4.21. Let $X = \mathbf{V}(z - x^2 - y^2)$. If we take $\phi(x) \in \mathbb{C}[X]$ for ϕ from 4.1, then

$$Y = \mathbf{V}_X(\phi(x)) = \{(0, a, a^2) \mid a \in \mathbb{C}\} \subset X$$

is a subvariety of X. Note that this is the same as $\mathbf{V}(z - x^2 - y^2, x)$ in $\mathbb{C}^3$.

Proposition 4.22. *Let $X \subset \mathbb{C}^n$ be an affine variety. If $Y = \mathbf{V}_X(f_1, \ldots, f_m)$ is a subvariety of X, then $Y = \mathbf{V}_X(\mathbf{I}_X(Y))$.*

Proof. Since every $f \in \mathbf{I}_X(Y)$ vanishes on Y, the inclusion $Y \subseteq \mathbf{V}_X(\mathbf{I}_X(Y))$ follows directly from the definition of $\mathbf{V}_X(J)$ for an ideal J. Going the other way, note that $f_1, \ldots, f_m \in \mathbf{I}_X(Y)$ by the definition of $\mathbf{I}_X$. It follows that $\mathbf{V}_X(\mathbf{I}_X(Y)) \subseteq \mathbf{V}_X(f_1, \ldots, f_m) = Y$ because if every function in $\mathbf{I}_X(Y)$ vanishes on some $p \in X$, then $f_1, \ldots, f_m$ vanish on p. □

Proposition 4.23. *Let $X \subset \mathbb{C}^n$ be an affine variety. Then every ideal $J \subset \mathbb{C}[X]$ is finitely generated.*

Proof. Let $I = \phi^{-1}(J) \subseteq \mathbb{C}[x_1, \ldots, x_n]$, where ϕ is the ring homomorphism (4.1). We claim that I is an ideal. Indeed, $0 \in I$ since $\phi(0) = 0 \in J$. Take $F, G \in I$ and $H \in \mathbb{C}[x_1, \ldots, x_n]$. Then

$$\phi(F+G) = \phi(F) + \phi(G) \in J \Rightarrow F + G \in \phi^{-1}(J),$$
$$\phi(HF) = \phi(H)\phi(F) \in J \Rightarrow HF \in \phi^{-1}(J).$$

By Proposition 4.6, $I = \langle F_1, \ldots, F_m \rangle$ for some $F_1, \ldots, F_m \in \mathbb{C}[x_1, \ldots, x_n]$. We claim that

$$J = \langle \phi(F_1), \ldots, \phi(F_m) \rangle.$$

The inclusion $\supseteq$ is obvious since $\phi(F_1), \ldots, \phi(F_m) \in J$. For the reverse inclusion, let $f \in J$. Take some $F \in \phi^{-1}(f) \subseteq I$. Since $I = \langle F_1, \ldots, F_m \rangle$, then $F = \sum_{i=1}^{m} G_i F_i$ for some $G_i \in \mathbb{C}[x_1, \ldots, x_n]$. Hence

$$f = \phi(F) = \sum_{i=1}^{m} \phi(G_i)\phi(F_i) \in \langle \phi(F_1), \ldots, \phi(F_m) \rangle.$$

□

If we let $X = \mathbb{C}^n$ then every affine variety Y is a subvariety of X. This is because X is an affine variety: it is defined by the zero polynomial $0 \in \mathbb{C}[x_1, \ldots, x_n]$. Hence if $Y = \mathbf{V}(F_1, \ldots, F_m)$ for $F_1, \ldots, F_m \in \mathbb{C}[x_1, \ldots, x_n]$, then $Y = \mathbf{V}_X(\phi(F_1), \ldots, \phi(F_m))$. Thus, everything described in this section can be applied to affine varieties too.

4.4 Zariski topology

It is crucial that every affine variety can be turned into a topological space. But before giving the proof of this we should be familiar with the concepts of sum and product of ideals.

Definition 4.24. Let $\{J_\alpha\}_{\alpha \in A}$ be a (possibly infinite) collection of ideals of a ring R. We define the sum of J_α to be

$$\sum_{\alpha \in A} J_\alpha = \left\{ \sum_{i=1}^{m} f_i \;\middle|\; f_i \in J_{\alpha_i} \text{ for some } \alpha_i \in A, m \in \mathbb{N} \right\}.$$

In other words, this is the set of all finite sums of elements from $\{J_\alpha\}_{\alpha \in A}$.

Proposition 4.25. *Let $\{J_\alpha\}_{\alpha \in A}$ be an arbitrary collection of ideals of R. Then $\sum_{\alpha \in A} J_\alpha$ is an ideal of R.*

Proof. Obviously 0 belongs to the sum since it belongs to each of these ideals. Take $f, g \in \sum_{\alpha \in A} J_\alpha$ and $h \in R$. Then $f = \sum_{i=1}^{m_1} f_i$ and $g = \sum_{i=1}^{m_2} g_i$ are finite sums of elements from $\{J_\alpha\}_{\alpha \in A}$. Hence

$$f + g = \sum_{i=1}^{m_1} f_i + \sum_{i=1}^{m_2} g_i \in \sum_{\alpha \in A} J_\alpha,$$
$$hf = h \sum_{i=1}^{m_1} f_i = \sum_{i=1}^{m_1} h f_i \in \sum_{\alpha \in A} J_\alpha.$$

□

Proposition 4.26. *Let X C^n be an a ne variety and fJ_αg$_{\alpha 2 A}$ be an arbitrary collection of ideals of* C$[X]$. *Then*

$$\bigcap_{\alpha 2 A} \mathbf{V}_X(J_\alpha) = \mathbf{V}_X\left(\sum_{\alpha 2 A} J_\alpha\right).$$

Proof. : Take p *2* $\bigcap_{\alpha 2 A} \mathbf{V}_X(J_\alpha)$ and f *2* $\sum_{\alpha 2 A} J_\alpha$. Then $f = \sum_{i=1}^{m_1} f_i$ is a finite sum of elements from fJ_αg$_{\alpha 2 A}$. Since f_i belongs to J_{α_i} for some α_i *2* A, then $f_i(p) = 0$. Hence $f(p) = \sum_{i=1}^{m_1} f_i(p) = 0$.

: Take p *2* $\mathbf{V}_X\left(\sum_{\alpha 2 A} J_\alpha\right)$. Since J_α *2* $\sum_{\alpha 2 A} J_\alpha$ for every α *2* A, then p *2* $\mathbf{V}_X(J_\alpha)$ for every α *2* A. Hence p *2* $\bigcap_{\alpha 2 A} \mathbf{V}_X(J_\alpha)$. □

We can ask a question: is an arbitrary union of subvarieties of X also a subvariety of X? It turns out that the answer is no in general.

Example 4.27. Let $Y = \bigcup_{a 2 \mathsf{Z}} \mathbf{V}_{\mathsf{C}}(x \quad a)$ be an infinite subset of C. Then Y is not a subvariety of C since every polynomial in C$[x]$ has only finitely many roots.

However, the answer to the above question becomes positive for a finite union of subvarieties. Notice that if we prove that for a union of two subvarieties, then by induction we can extend the proof to any finite union. It turns out that the union of two subvarieties corresponds algebraically to the notion of product of two ideals.

Definition 4.28. Let I and J be two ideals of a ring R. We define the product of I and J to be

$$I \quad J = \left\{ \sum_{i=1}^{m} f_i g_i \; \mathsf{j} \; f_1, \ldots, f_m \; 2 \; I, g_1, \ldots, g_m \; 2 \; J, m \; 2 \; \mathsf{N} \right\}.$$

It is easy to see that $I \quad J$ is an ideal of R. Obviously 0 *2* $I \quad J$ as it belongs to both I and J. For f, g *2* $I \quad J$ and h *2* R we have $f + g$ *2* $I \quad J$ and hf *2* $I \quad J$.

Proposition 4.29. *Let X C^n be an a ne variety and I, J be two ideals of* C$[X]$. *Then*

$$\mathbf{V}_X(I) \; [\; \mathbf{V}_X(J) = \mathbf{V}_X(I \quad J).$$

Proof. : Take p *2* $\mathbf{V}_X(I) \; [\; \mathbf{V}_X(J)$. Then either $f(p) = 0$ for all f *2* I or $g(p) = 0$ for all g *2* J. Thus, $f(p)g(p) = 0$ for all f *2* I and all g *2* J. Thus, $h(p) = 0$ for all h *2* $I \quad J$ and, hence, p *2* $\mathbf{V}_X(I \quad J)$.

: Take p *2* $\mathbf{V}_X(I \quad J)$. Then $f(p)g(p) = 0$ for all f *2* I and all g *2* J. If $f(p) = 0$ for all f *2* I, then p *2* $\mathbf{V}_X(I)$. If $f(p) \neq 0$ for some f *2* I, then we must have $g(p) = 0$ for all g *2* J. In either case, p *2* $\mathbf{V}_X(I) \; [\; \mathbf{V}_X(J)$. □

Now we are ready to define a topology on an a□ne variety X (see Definition 2.28).

Proposition 4.30. *Let X C^n be an a ne variety. Let τ be the set of all subvarieties of X. Then (X, τ) is a topological space.*

Proof. 1. ? and X belong to τ since

$$? = \mathbf{V}_X(1), \quad X = \mathbf{V}_X(0).$$

2. Let $\{Y_\alpha\}_{\alpha \in A}$ be an arbitrary collection of subvarieties of X. Denote by J_α the ideal generated by the functions which define Y_α. By Proposition 4.20, $Y_\alpha = \mathbf{V}_X(J_\alpha)$. By Proposition 4.26,

$$\bigcap_{\alpha \in A} Y_\alpha = \mathbf{V}_X\left(\sum_{\alpha \in A} J_\alpha\right).$$

According to Proposition 4.23, $\sum_{\alpha \in A} J_\alpha = \langle f_1, \ldots, f_m \rangle$ for some $f_1, \ldots, f_m \in \mathbb{C}[X]$. Hence, by Proposition 4.20,

$$\bigcap_{\alpha \in A} Y_\alpha = \mathbf{V}_X(f_1, \ldots, f_m),$$

which means $\bigcap_{\alpha \in A} Y_\alpha$ is a subvariety of X.

3. We need to prove that any finite union of subvarieties of X is again a subvariety of X. However, this is equivalent to proving the same statement for a union of two subvarieties. So, let Y_1, Y_2 be subvarieties of X. Denote J_1 and J_2 the ideals generated by the functions which define Y_1 and Y_2, respectively. By Proposition 4.20, $Y_i = \mathbf{V}_X(J_i)$ for $i = 1, 2$. By Proposition 4.29,

$$Y_1 \cup Y_2 = \mathbf{V}_X(J_1 \cap J_2).$$

According to Proposition 4.23, $J_1 \cap J_2 = \langle g_1, \ldots, g_r \rangle$ for some $g_1, \ldots, g_r \in \mathbb{C}[X]$. Hence, by Proposition 4.20,

$$Y_1 \cup Y_2 = \mathbf{V}_X(g_1, \ldots, g_r),$$

which means $Y_1 \cup Y_2$ is a subvariety of X.

□

The topology defined in Proposition 4.30 is called the Zariski topology on X.

4.5 Rational maps

Let $X \subseteq \mathbb{C}^n$ and $Y \subseteq \mathbb{C}^m$ be irreducible affine varieties. Hence we may talk about function fields $\mathbb{C}(X)$ and $\mathbb{C}(Y)$. We define the notion of rational map $\varphi \colon X \dashrightarrow Y$.

Definition 4.31. A rational map $\varphi \colon X \dashrightarrow Y$ is an m-tuple of rational functions $\varphi_1, \ldots, \varphi_m \in \mathbb{C}(X)$ such that, for all points $p \in X$ at which all the φ_i are regular, $\varphi(p) \stackrel{\text{def}}{=} (\varphi_1(p), \ldots, \varphi_m(p)) \in Y$. We say that φ is regular at such a point p and $\varphi(p) \in Y$ is the image of p under φ. The image of X under a rational map φ is the set of points

$$\varphi(X) \stackrel{\text{def}}{=} \{\varphi(p) \mid p \in X \text{ and } \varphi \text{ is regular at } p\}.$$

The set of all points at which φ is regular is called the domain of φ.

Proposition 4.32. *Let φ 2 C(X) be a rational function. Then the set of points p 2 X at which φ is regular is nonempty and open in X.*

Proof. Take any representative $\frac{f}{g}$ of φ. Because g 6= 0 2 C$[X]$, then this exactly means that $g(p)$ 6= 0 for some p 2 X. So, φ is regular at p. To prove that the set of such points is open, consider all possible representations $\varphi = \frac{f_i}{g_i}$. For any regular function g_i the set Y_i X of points p 2 X for which $g_i(p) = 0$ is obviously closed, and hence $U_i = X$ n Y_i is open. The set U of points at which φ is regular is by definition $U = \bigcup U_i$, and therefore is open. By Proposition 2.44, U is dense in X. □

Remark 4.33. For any rational map φ: X 99K Y the set of points p 2 X at which φ is regular is nonempty and open in X. This follows directly from Proposition 2.46. Again, by Proposition 2.44, U is dense in X.

Example 4.34. Let $X = \mathbf{V}(y \quad x^2)$ to $Y =$ C. We define the map

$$\begin{aligned} \varphi\colon X &\text{ 99K } Y \\ (x, y) &\text{ 7! } x \end{aligned}$$

This is a rational map since, using the notation of Example 4.17, it is defined by a rational function $\left[\frac{x}{1}\right]$ 2 C(X). Thus, φ is defined everywhere on X.

Example 4.35. Let $X =$ C to $Y = \mathbf{V}(xy \quad 1)$. We define the map

$$\begin{aligned} \varphi\colon X &\text{ 99K } Y \\ t &\text{ 7! } \left(t, \frac{1}{t}\right) \end{aligned}$$

This is a rational map since it is defined by rational functions $\left[\frac{t}{1}\right], \left[\frac{1}{t}\right]$ 2 C(X). We note that φ is defined on C$nf0g$ which is nonempty and open and, hence, dense in C.

Proposition 4.36. *Let φ 2 C(X) be a rational function. If φ vanishes on some nonempty open subset U X, then $\varphi = 0$ 2 C(X).*

Proof. Take a representative $\frac{f}{g}$ of φ. We need to prove that $f = 0$ 2 C$[X]$, or that $\mathbf{V}_X(f) = X$. The inclusion $\mathbf{V}_X(f)$ X is obvious. For the reverse inclusion let $Y = X$ n $\mathbf{V}_X(g)$ be an open subset of X. Then f vanishes on the nonempty open subset $U \setminus Y$ X. By Proposition 2.44, $\overline{U \setminus Y} = X$. By Corollary 2.32, $X = \overline{U \setminus Y}$ $\mathbf{V}_X(f)$. □

We now find out how a rational map φ: X 99K Y induces a map on C$[Y]$. We define

$$\begin{aligned} \varphi^{\#}\colon \mathrm{C}[Y] &\text{ ! } \mathrm{C}(X) \\ g &\text{ 7! } G(\varphi_1, \ldots, \varphi_m), \end{aligned}$$

where G 2 C$[x_1, \ldots, x_m]$ is a polynomial defining g. We can understand $G(\varphi_1, \ldots, \varphi_m)$ as an element of C(X): we can add and multiply $\varphi_1, \ldots, \varphi_m$ inside C(X) provided we understand the coe□cients of G as constant elements of C(X).

Proposition 4.37. *Let φ: X 99K Y be a rational map. Then:*

1. *$\varphi^{\#}$ is well-defined.*

2. *$\varphi^{\#}$ is a ring homomorphism.*

Proof. 1. Let $G, G' \in \mathsf{C}[x_1, \ldots, x_m]$ be two different polynomials which represent g. We need to show that

$$G(\varphi_1, \ldots, \varphi_m) = G'(\varphi_1, \ldots, \varphi_m),$$

or, equivalently,

$$(G - G')(\varphi_1, \ldots, \varphi_m) = 0 \in \mathsf{C}(X).$$

We know that $G - G' \in \mathbf{I}(Y)$. If we show that $H(\varphi_1, \ldots, \varphi_m) = 0 \in \mathsf{C}(X)$ for any $H \in \mathbf{I}(Y)$, then we are done. So, take any $H \in \mathbf{I}(Y)$. Then, by Remark 4.33, $H(\varphi_1, \ldots, \varphi_m) \in \mathsf{C}(X)$ is defined on some nonempty open subset U of X. Since $(\varphi_1(p), \ldots, \varphi_m(p)) \in Y$ for any $p \in U$, then $(H(\varphi_1, \ldots, \varphi_m))(p) = H(\varphi_1(p), \ldots, \varphi_m(p)) = 0$ for all $p \in U$. By Proposition 4.36, $H(\varphi_1, \ldots, \varphi_m) = 0 \in \mathsf{C}(X)$.

2. It is obvious that $\varphi^{\#}(1_{\mathsf{C}[Y]}) = 1_{\mathsf{C}(X)}$. We should check that $\varphi^{\#}(g_1 + g_2) = \varphi^{\#}(g_1) + \varphi^{\#}(g_2)$ and $\varphi^{\#}(g_1 \cdot g_2) = \varphi^{\#}(g_1) \cdot \varphi^{\#}(g_2)$. Let G_1 and G_2 be polynomials which represent g_1 and g_2, respectively. Hence $G_1 + G_2$ represents $g_1 + g_2$. Then,

$$\varphi^{\#}(g_1+g_2) = (G_1+G_2)(\varphi_1, \ldots, \varphi_m) = G_1(\varphi_1, \ldots, \varphi_m)+G_2(\varphi_1, \ldots, \varphi_m) = \varphi^{\#}(g_1)+\varphi^{\#}(g_2).$$

The equality $\varphi^{\#}(g_1 \cdot g_2) = \varphi^{\#}(g_1) \cdot \varphi^{\#}(g_2)$ can be verified in exactly the same way.

□

It will be useful for us to understand when a homomorphism $\varphi^{\#} : \mathsf{C}[Y] \to \mathsf{C}(X)$ corresponding to a regular map $\varphi : X \dashrightarrow Y$ is injective. We give the following proposition.

Proposition 4.38. *Let $\varphi : X \dashrightarrow Y$ be a rational map. Then a homomorphism $\varphi^{\#} : \mathsf{C}[Y] \to \mathsf{C}(X)$ is injective if and only if $\varphi(X)$ is dense in Y.*

Proof. By Proposition 2.15, $\varphi^{\#}$ is injective if and only if $\ker(\varphi^{\#})$ is trivial. Take $g \in \ker(\varphi^{\#})$ and let U be the domain of φ which is, by Remark 4.33, nonempty and open. Then $g(\varphi(p)) = 0$ for all $p \in U$, or, equivalently, g vanishes on $\varphi(X)$. But because $\mathbf{V}_Y(g)$ is a closed set in Y, then, by Corollary 2.32, $\overline{\varphi(X)} \subseteq \mathbf{V}_Y(g)$. Hence, $g \in \mathbf{I}_Y(\overline{\varphi(X)})$. It is also easy to see that $\mathbf{I}_Y(\overline{\varphi(X)}) \subseteq \ker(\varphi^{\#})$. Indeed, $g \in \mathbf{I}_Y(\overline{\varphi(X)})$ implies g vanishes on $\varphi(X)$ since $\varphi(X) \subseteq \overline{\varphi(X)}$. Then $\varphi^{\#}(g) = G(\varphi_1, \ldots, \varphi_m) \in \mathsf{C}(X)$ vanishes on U. By Proposition 4.36, $\varphi^{\#}(g) = 0 \in \mathsf{C}(X)$. Thus, $\mathbf{I}_Y(\overline{\varphi(X)}) = \ker(\varphi^{\#})$. It follows that the triviality of $\ker(\varphi^{\#})$ is equivalent to the triviality of $\mathbf{I}_Y(\overline{\varphi(X)})$. And the triviality of $\mathbf{I}_Y(\overline{\varphi(X)})$ is equivalent to $\overline{\varphi(X)} = Y$ because, using Proposition 4.22,

$$\mathbf{I}_Y(\overline{\varphi(X)}) = \{0\} \Rightarrow \overline{\varphi(X)} = \mathbf{V}_Y(\mathbf{I}_Y(\overline{\varphi(X)})) = \mathbf{V}_Y(\{0\}) = Y,$$

$$\overline{\varphi(X)} = Y \Rightarrow \mathbf{I}_Y(\overline{\varphi(X)}) = \mathbf{I}_Y(Y) = \{0\}.$$

□

Example 4.39. Recall Example 4.35. For $X = \mathbb{C}$ to $Y = \mathbf{V}(xy \quad 1)$ we define a rational map

$$\begin{aligned} \varphi\colon X &\dashrightarrow Y \\ t &\mapsto \left(t, \frac{1}{t}\right) \end{aligned}$$

Its image is $\varphi(X) = Y$. Hence, by Proposition 4.38, $\varphi^{\#}\colon \mathbb{C}[Y] \quad ! \quad \mathbb{C}(X)$ is injective. An advanced reader may notice that we can interpret $\varphi^{\#}$ as a ring embedding $\mathbb{C}[x, x \quad ^{1}] \hookrightarrow \mathbb{C}(x)$, where $\mathbb{C}[x, x \quad ^{1}]$ is the localization of $\mathbb{C}[x]$ at x.

We can define a rational map from Y to X in the following way

$$\begin{aligned} \psi\colon Y &\dashrightarrow X \\ (x, y) &\mapsto x \end{aligned}$$

Its image is $\psi(Y) = \mathbb{C}nf0g$. We know that in the Zariski topology the closed subsets of $\mathbb{C}$ are the empty set, finite subsets of $\mathbb{C}$ and the whole $\mathbb{C}$. This is because every nonconstant polynomial in $\mathbb{C}[x]$ has finitely many roots. That's why the closure of $\mathbb{C}nf0g$ in $\mathbb{C}$ is the whole $\mathbb{C}$, which, by Proposition 4.38, means that $\psi^{\#}\colon \mathbb{C}[X] \quad ! \quad \mathbb{C}(Y)$ is injective. Then, $\mathbb{C}(Y) = \mathrm{Frac}(\mathbb{C}[x, x \quad ^{1}]) = \mathbb{C}(x)$. Then we can interpret $\psi^{\#}$ as a ring embedding $\mathbb{C}[x] \hookrightarrow \mathbb{C}(x)$.

Example 4.40. Let $X = \mathbf{V}(x + y \quad 1)$ and $Y = \mathbb{C}$. Then we define a rational map

$$\begin{aligned} \varphi\colon X &\dashrightarrow Y \\ (x, y) &\mapsto \frac{1}{x + y} \end{aligned}$$

It is obvious that $\varphi(X) = f1g$. The closure $\overline{\varphi(X)}$ of $\varphi(X)$ in Y is equal to $\varphi(X)$ since $\varphi(X)$ is finite. By Proposition 4.38, $\varphi^{\#}\colon \mathbb{C}[Y] \quad ! \quad \mathbb{C}(X)$ is not injective. Since $\mathbb{C}[Y] = \mathbb{C}[\mathbb{C}]$ may be identified with $\mathbb{C}[t]$ using the isomorphism (4.2), then $\ker(\varphi^{\#}) = ht \quad 1i \quad \mathbb{C}[t]$.

Definition 4.41. A rational map $\varphi\colon X \dashrightarrow Y$ with $\varphi(X)$ is called dominant if $\varphi(X)$ is dense in Y.

The next proposition shows that for a dominant rational map $\varphi\colon X \dashrightarrow Y$ we may construct a map $\varphi \quad : \mathbb{C}(Y) \quad ! \quad \mathbb{C}(X)$.

Proposition 4.42. *Let $\varphi\colon X \dashrightarrow Y$ be a dominant rational map. Then there is an injective ring homomorphism $\varphi \quad : \mathbb{C}(Y) \quad ! \quad \mathbb{C}(X)$ of function fields. We call φ the inclusion of $\mathbb{C}(Y)$ into $\mathbb{C}(X)$ and write $\varphi \quad : \mathbb{C}(Y) \hookrightarrow \mathbb{C}(X)$.*

Proof. According to Proposition 2.24, we define φ to be an extension of an injective ring homomorphism $\varphi^{\#}\colon \mathbb{C}[Y] \quad ! \quad \mathbb{C}(X)$. □

Notice that the inclusion φ sends constant functions $\mathbb{C} \quad \mathbb{C}(Y)$ to the same constant functions $\mathbb{C} \quad \mathbb{C}(X)$. We say that φ is the identity on constants. So, Proposition 4.42 tells us that any dominant rational map induces an inclusion of function fields. It can be shown that the converse is also true: any inclusion of function fields gives rise to a dominant rational map. In fact, there is a bijection between these sets.

Proposition 4.43. *Let X and Y be irreducible a ne varieties. Then the operator $\varphi \mapsto \varphi$ yields a bijection between dominant rational maps X 99K Y and injective ring homomorphisms $\mathsf{C}(Y) \to \mathsf{C}(X)$ which are identities on constants.*

Proof. [26, Theorem 16]. □

Let $X \subseteq \mathsf{C}^n, Y \subseteq \mathsf{C}^m, Z \subseteq \mathsf{C}^r$ be three irreducible a□ne varieties. Given two rational maps $\varphi\colon X$ 99K Y and $\psi\colon Y$ 99K Z such that φ is dominant, it is easy to see that we can define a composite $\psi \circ \varphi\colon X$ 99K Z as follows. Let ψ be given by $\psi_1 = \left[\frac{f_1}{g_1}\right], \ldots, \psi_r = \left[\frac{f_r}{g_r}\right] \in \mathsf{C}(Y)$ with $f_i, g_i \in \mathsf{C}[Y]$. If φ is defined by $\varphi_1, \ldots, \varphi_m \in \mathsf{C}(X)$ then we define

$$\psi \circ \varphi \stackrel{\text{def}}{=} \left(\frac{F_1(\varphi_1, \ldots, \varphi_m)}{G_1(\varphi_1, \ldots, \varphi_m)}, \ldots, \frac{F_r(\varphi_1, \ldots, \varphi_m)}{G_r(\varphi_1, \ldots, \varphi_m)}\right),$$

where $F_i, G_i \in \mathsf{C}[x_1, \ldots, x_m]$ are polynomials defining f_i, g_i. It can be verified that $\psi \circ \varphi$ is well-defined, i.e. the result doesn't depend on the choice of representatives F_i, G_i and the denominators are not the zero functions on X. If in addition ψ is dominant then so is $\psi \circ \varphi$.

Proposition 4.44. *If $\varphi\colon X$ 99K Y and $\psi\colon Y$ 99K Z are dominant rational maps, then $(\psi \circ \varphi) = \varphi \circ \psi$.*

Proof. [24, Section 3.3]. □

Definition 4.45. A dominant rational map $\varphi\colon X$ 99K Y is said to be birational if there is a dominant rational map $\psi\colon Y$ 99K X such that $\psi \circ \varphi = \mathrm{id}_X, \varphi \circ \psi = \mathrm{id}_Y$ (where defined).

Proposition 4.46. *Let X and Y be irreducible a ne varieties. Then the operator $\varphi \mapsto \varphi$ yields a bijection between birational maps X 99K Y and ring isomorphisms $\mathsf{C}(Y) \to \mathsf{C}(X)$ which are identities on constants.*

Proof. [26, Corollary 18] or [6, Chapter 5, §5, Theorem 10]. □

5 Elements of Algebraic Topology

Algebraic topology is a branch of science which studies topological spaces from algebraic perspective, i.e. to every topological space it associates an algebraic object, and to every continuous map between topological spaces it associates a homomorphism between their associated algebraic objects. For example, in Chapter 4 we saw that to every a□ne variety (with the Zariski topology on it) we can associate an algebraic object called the function field so that a dominant rational map $f : X$ 99K Y between a□ne varieties (which is continuous in the Zariski topology) induces a ring homomorphism of function fields $f : \mathrm{C}(Y) \circ ! \mathrm{C}(X)$. Algebraic topology studies general topological spaces. These can also be a□ne varieties with the topology induced from the standard topology of C^n (we will explain it in detail in Chapter 7).

In this chapter we are going to show that to every topological space we can associate an algebraic object, called the fundamental group. This algebraic object is constructed from the topological space by forming the loops in that space, i.e. the paths starting and ending at the same point. We also show how the fundamental group is connected to a special kind of maps between topological spaces, called covering maps. Finally, we explain how the symmetries of the covering map can be revealed using the fundamental group.

We recommend [15] as a good introduction to algebraic topology. It's very well written and it isn't supposed that the reader already knows a lot from algebra and general topology.

5.1 Fundamental group

Before talking about the fundamental group we have to be familiar with the concept of a path in a space. Recall the definition of a continuous map between topological spaces (Definition 2.35).

Definition 5.1. A path in a topological space X is a continuous map $f : I ! X$ where I is the unit interval $[0, 1]$ R.

Example 5.2. Let X be the unit circle in R^2 with the topology induced from R^2. We define a map

$$\begin{aligned} f : I &! X \\ s &7! (\cos(s \ \pi), \sin(s \ \pi)) \end{aligned}$$

It can be verified that f is continuous. So, f is a path in X from 0 to π radians.

We will be interested in deformations of paths.

Definition 5.3. A homotopy of paths in a topological space X is a family $f_t\colon I \to X$, $0 \le t \le 1$ of paths, such that the map $F\colon I \times I \to X$ defined by $F(s,t) = f_t(s)$ is continuous. When two paths f_0 and f_1 are connected in this way by a homotopy f_t, they are said to be homotopic. We will also say that f_t is a homotopy between f_0 and f_1, or, simply, f_t is a homotopy of f_0.

Example 5.4. Let $X = \mathbb{R}$. Let also f_0 be a path in X which sends all $s \in I$ to $0 \in X$ (so, f_0 is just a constant path) and f_1 be a path in X which sends $s \in I$ to $s \in X$ (f_1 defines a unit interval in X). Define

$$\begin{aligned} f_t\colon I &\to X \\ s &\mapsto t - s \end{aligned}$$

It is obvious that F defined by $F(s,t) = f_t(s)$ is continuous. Thus, f_t is a homotopy between f_0 and f_1.

It can be observed that a homotopy defined in Example 5.4 doesn't fix the endpoints of f_0 and f_1, because $f_t(1)$ changes as t varies. In order to define the fundamental group of a topological space we need a little bit different concept.

Definition 5.5. A homotopy of paths in X, fixing the endpoints, is a homotopy f_t in X such that

$$f_t(0) = x_0 \text{ and } f_t(1) = x_1 \quad \text{for all } 0 \le t \le 1,$$

i.e. the endpoints are the same for all t. If such a homotopy between f_0 and f_1 exists, then we say that f_0 and f_1 are homotopic relatively to the endpoints (or, simply, homotopic r.t.e.).

Example 5.6. Let $X = \mathbb{R}^2$. Let f_0 be the path which sends $s \in I$ to $(1-2s, 0) \in X$ and f_1 be a path from Example 5.2 (see Figure 5.1). Then we define

$$\begin{aligned} f_t\colon I &\to X \\ s &\mapsto (1-t)f_0(s) + tf_1(s) = \Big((1-t)(1-2s) + t\cos(s-\pi), t\sin(s-\pi)\Big) \end{aligned}$$

This is a homotopy between f_0 and f_1 since $F(s,t) = f_t(s)$ is continuous. Also f_t is a homotopy, fixing the endpoints, since $F(0,t) = \big((1-t)\cdot 1 + t\cos(0), t\sin(0)\big) = (1,0)$ and $F(1,t) = \big((1-t)\cdot(-1) + t\cos(\pi), t\sin(\pi)\big) = (-1,0)$ for all $0 \le t \le 1$.

So, given two paths f_0, f_1 in X with the same endpoints (i.e. $f_0(0) = f_1(0)$ and $f_0(1) = f_1(1)$) we can ask if they are homotopic r.t.e. Before proceeding further we give the following proposition.

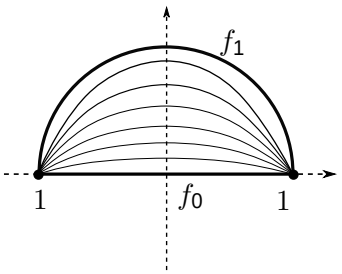


Figure 5.1

Proposition 5.7. *The relation of homotopy of paths with fixed endpoints in any space is an equivalence relation.*

Proof. [15, Proposition 1.2]. □

Given two paths $f, g\colon I \to X$ such that $f(1) = g(0)$, there is a composition path $f \cdot g$ that traverses first f and then g, defined by the formula

$$f \cdot g(s) \stackrel{\text{def}}{=} \begin{cases} f(2s), & 0 \le s \le \frac{1}{2} \\ g(2s-1), & \frac{1}{2} \le s \le 1 \end{cases}$$

Thus f and g are traversed twice as fast in order for $f \cdot g$ to be traversed in unit time.

Suppose we restrict attention to paths $f\colon I \to X$ with the same starting and ending point $f(0) = f(1) = x \in X$. Such paths are called loops, and the common starting and ending point x is called the basepoint. Then, according to Proposition 5.7, we can form the set of all homotopy classes $[f]$ of loops in X. This set is denoted $\pi_1(X, x)$.

Proposition 5.8. *$\pi_1(X, x)$ is a group with respect to the operation defined as $[f][g] \stackrel{\text{def}}{=} [f \cdot g]$.*

Proof. [15, Proposition 1.3]. □

This group is called the fundamental group of X at the basepoint x. For example, the fundamental group of the point is trivial, since there is a unique path from the unit interval to a point. It can be proven (see [15, Theorem 1.7]) that the fundamental group of the circle is isomorphic to Z (it reflects the fact that a positive integer n is represented by making n loops clockwise, while a negative integer n is represented by making n loops counterclockwise). In Figure 5.2 there is a space X with a black hole. The loop f is homotopic to a constant loop based at x. The loop g is not homotopic to this constant loop, since there is a hole inside this loop which forbids it to be deformed to a point. It can be shown (in the same way as for the circle) that the fundamental group of this space is $\pi_1(X, x) = \mathbb{Z}$.

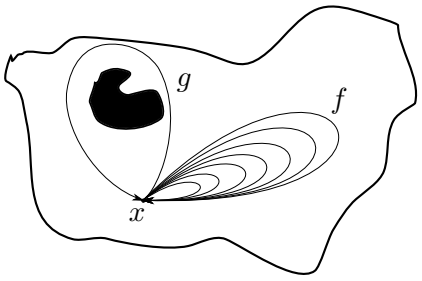


Figure 5.2

Given a path $f\colon I \to X$ we define the inverse path $\bar{f} : I \to X$ to be

$$\bar{f}(s) \stackrel{\text{def}}{=} f(1-s).$$

If f is a loop in X based at x, then $[\bar{f}] = [f]^{-1} \in \pi_1(X, x)$ [15, p. 26].

Suppose $\varphi\colon X \to Y$ is a map between topological spaces taking $x \in X$ to $y \in Y$. Then φ induces a homomorphism $\varphi_* : \pi_1(X, x) \to \pi_1(Y, y)$, defined by composing loops $f\colon I \to X$ based at x with φ, that is $\varphi_*([f]) = [\varphi f]$. This induced map φ_* is well-defined wince a homotopy f_t of loops based at x yields a composed $\varphi_*([f_0]) = [\varphi f_0] = [\varphi f_1] = \varphi_*([f_1])$. Moreover, φ_* is a homomorphism since $\varphi(f \cdot g) = (\varphi f) \cdot (\varphi g)$, both functions having the value $\varphi f(2s)$ for $0 \le s \le \frac{1}{2}$ and the value $\varphi g(2s-1)$ for $\frac{1}{2} \le s \le 1$.

It is also true that $(\varphi\psi)_* = \varphi_* \psi_*$. This follows from the fact that composition of maps is associative, so $(\varphi\psi)f = \varphi(\psi f)$.

Finally, we would like to give a general definition of a homotopy which we will use in Section 5.2.

Definition 5.9. A homotopy in a topological space X is a family $f_t : Y \to X$, $0 \le t \le 1$ of continuous functions, such that the map $F : Y \times I \to X$ defined by $F(y,t) = f_t(y)$ is continuous. When two functions f_0 and f_1 are connected in this way by a homotopy f_t, they are said to be homotopic and we write $f_0 \simeq f_1$. We will also say that f_t is a homotopy between f_0 and f_1, or, simply, f_t is a homotopy of f_0.

5.2 Covering spaces

Definition 5.10. Let X be a topological space. A covering space of X is a topological space $\widetilde{X}$ together with a continuous map $p : \widetilde{X} \to X$ such that for every $x \in X$ there exists a neighbourhood $x \in U \subset X$, such that $p^{-1}(U)$ is a union of disjoint open sets in $\widetilde{X}$, each of which is mapped homeomorphically onto U by p. The map p is called the covering map.

For example, let X be the unit circle in $\mathbb{R}^2$ and $\widetilde{X}$ be the image of the map

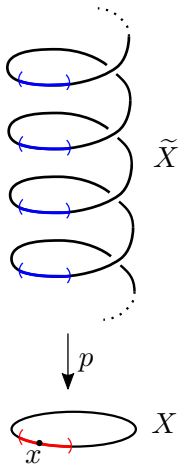


Figure 5.3

$$f : \mathbb{R} \to \mathbb{R}^3$$
$$s \mapsto (\cos(2\pi s), \sin(2\pi s), s)$$

which is just the spiral in $\mathbb{R}^3$ (see Figure 5.3). We define a map

$$p : \widetilde{X} \to X$$
$$(\cos(2\pi s), \sin(2\pi s), s) \mapsto (\cos(2\pi s), \sin(2\pi s))$$

Take $x = (\cos(2\pi s), \sin(2\pi s)) \in X$. Then the preimage $p^{-1}(x)$ is the set

$$p^{-1}(x) = \{(\cos(2\pi s), \sin(2\pi s), s + k) \mid k \in \mathbb{Z}\}.$$

Take an open neighbourhood $x \in U \subset X$ (red in Figure 5.3). The preimage $p^{-1}(U)$ is a union of disjoint open subsets of $\widetilde{X}$ (blue in Figure 5.3). Thus, $p : \widetilde{X} \to X$ is a covering space.

Remark 5.11. Let $p : \widetilde{X} \to X$ be a covering space and $x \in X$. Suppose $p^{-1}(x)$ is finite and that X is connected. Denote $d = |p^{-1}(x)|$. If we take an open neighbourhood $x \in U \subset X$ from Definition 5.10, then $|p^{-1}(u)| = d$ for all $u \in U$. As we vary x in X we see that the number d is locally constant, so it is constant on the whole X, i.e. $|p^{-1}(x)| = d$ for all $x \in X$. Such a covering space is called finite-sheeted.

We will be interested in lifts of paths (and their homotopies) in X under p. We give the following definition.

Definition 5.12. Let $p : \widetilde{X} \to X$ be a covering space. A lift of a map $f : Y \to X$ is a map $\widetilde{f} : Y \to \widetilde{X}$ such that $p\widetilde{f} = f$.

Proposition 5.13. *Given a covering space $p : \widetilde{X} \to X$, a homotopy $f_t : Y \to X$, and a map $\widetilde{f}_0 : Y \to \widetilde{X}$ lifting f_0, then there exists a unique homotopy $\widetilde{f}_t : Y \to \widetilde{X}$ of $\widetilde{f}_0$ that lifts f_t.*

Proof. [15, Proposition 1.30]. □

Taking Y to be a point in Proposition 5.13 gives the path lifting property for a covering space $p: \widetilde{X} \to X$, which says (see Figure 5.4) that for each path $f: I \to X$ (green) and each lift $\widetilde{x}$ of the starting point $f(0) = x$ there is a unique path $\widetilde{f}: I \to \widetilde{X}$ starting at $\widetilde{x}$ that lifts f (red and blue). In particular, the uniqueness of lifts implies that every lift of a constant path is constant.

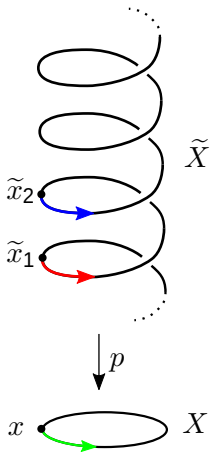


Figure 5.4

Remark 5.14. Taking Y to be I, we see that every homotopy f_t of a path f_0 in X, fixing the endpoints, lifts to a unique homotopy $\widetilde{f}_t$ of each lift $\widetilde{f}_0$ of f_0. The lifted homotopy $\widetilde{f}_t$ is also a homotopy of paths, fixing the endpoints, since as t varies each endpoint of $\widetilde{f}_t$ traces out a path lifting a constant path, which is constant as was noted in the previous paragraph.

Proposition 5.15. *The map $p_*: \pi_1(\widetilde{X}, \widetilde{x}) \to \pi_1(X, x)$ induced by a covering space $p: \widetilde{X} \to X$, taking $\widetilde{x}$ to x, is injective. The image subgroup $p_*(\pi_1(\widetilde{X}, \widetilde{x}))$ in $\pi_1(X, x)$ are exactly the homotopy classes of loops in X based at x whose lifts to $\widetilde{X}$ starting at $\widetilde{x}$ are loops.*

Proof. An element of the kernel of p_* is represented by a loop $\widetilde{f}_0: I \to \widetilde{X}$ such that there exists a homotopy $f_t: I \to X$ of $f_0 = p\widetilde{f}_0$ to the trivial loop f_1, fixing the endpoints. By Proposition 5.13, there is a lifted homotopy of loops $\widetilde{f}_t$ starting with $\widetilde{f}_0$ and ending with a constant loop. Hence $[\widetilde{f}_0] = e_{\pi_1(\widetilde{X}, \widetilde{x})}$ and p_* is injective.

For the second statement of the proposition, loops at x lifting to loops at $\widetilde{x}$ certainly represent elements of the image of $p_*: \pi_1(\widetilde{X}, \widetilde{x}) \to \pi_1(X, x)$. Conversely, a loop representing an element of the image of p_* is homotopic to a loop having such a lift, so by Proposition 5.13, the loop itself must have such a lift. □

The following proposition shows that for a path-connected covering space $p: \widetilde{X} \to X$ with $\widetilde{x}_1 \in p^{-1}(x)$, changing the basepoint $\widetilde{x}_1$ within $p^{-1}(x)$ corresponds exactly to changing $p_*(\pi_1(\widetilde{X}, \widetilde{x}_1))$ to a conjugate subgroup of $\pi_1(X, x)$.

Proposition 5.16. *Let $p: \widetilde{X} \to X$ be a path-connected covering space with $\widetilde{x}_1, \widetilde{x}_2 \in p^{-1}(x)$ and let H be the subgroup $p_*(\pi_1(\widetilde{X}, \widetilde{x}_1)) \subset \pi_1(X, x)$. Then $p_*(\pi_1(\widetilde{X}, \widetilde{x}_2)) = [\gamma]^{-1} H [\gamma]$ for some loop γ in X.*

Proof. Let $\widetilde{\gamma}$ be a path from $\widetilde{x}_1$ to $\widetilde{x}_2$. Then $\widetilde{\gamma}$ projects by p to a loop γ in X, representing an element $g = [\gamma] \in \pi_1(X, x)$. Set $H_i = p_*(\pi_1(\widetilde{X}, \widetilde{x}_i))$ for $i = 1, 2$. We have an inclusion $g^{-1} H_0 g \subset H_1$ since for a loop $\widetilde{\alpha}$ at $\widetilde{x}_1$, $\widetilde{\gamma} \cdot \widetilde{\alpha} \cdot \widetilde{\gamma}$ is a loop at $\widetilde{x}_2$. Similarly we have $g H_1 g^{-1} \subset H_0$. Conjugating the latter relation by g^{-1} gives $H_1 \subset g^{-1} H_0 g$, so $g^{-1} H_0 g = H_1$. Thus, changing the basepoint from $\widetilde{x}_1$ to $\widetilde{x}_2$ changes H_0 to the conjugate subgroup $H_1 = g^{-1} H_0 g$ of $\pi_1(X, x)$. □

5.3 Monodromy group

Let $p\colon \widetilde{X} \to X$ be a covering space. Take $x \in X$ and denote $F = p^{-1}(x)$. Then every loop γ based at x defines a function $\sigma_\gamma\colon F \to F$ in the following way:

$$\begin{aligned} \sigma_\gamma\colon F &\to F \\ \widetilde{x} &\mapsto \widetilde{\gamma}(1) \end{aligned}$$

where $\widetilde{\gamma}$ is a unique lift of γ starting at $\widetilde{x}$. We note that σ_γ acts bijectively on F since $\sigma_{\gamma^{-1}}$ is its inverse. By Remark 5.14, if two loops γ and γ' are homotopic r.t.e., then their lifts $\widetilde{\gamma}$ and $\widetilde{\gamma}'$ starting at $\widetilde{x}$ are homotopic r.t.e., i.e. $\widetilde{\gamma}(1) = \widetilde{\gamma}'(1)$. Thus, σ_γ depends only on the homotopy class of loops based at x. Hence we have a well-defined map:

$$\begin{aligned} \varphi\colon \pi_1(X,x) &\to \mathrm{S}(F) \\ [\gamma] &\mapsto \sigma_\gamma \end{aligned} \tag{5.1}$$

which sends a homotopy class of loops $[\gamma]$ to a permutation of the fiber F defined by γ. We show that φ is a group homomorphism. Let $[\gamma_1], [\gamma_2]$ be two homotopy classes of loops. Then

$$\varphi([\gamma_1][\gamma_2]) = \varphi([\gamma_1 \cdot \gamma_2]) = \sigma_{(\gamma_1 \cdot \gamma_2)} = \left(\widetilde{x} \mapsto (\gamma_1 \cdot \gamma_2)(1)\right) = \left(\widetilde{x} \mapsto \gamma_2 \cdot \gamma_1(1)\right) =$$

$$= \sigma_{\gamma_1}\sigma_{\gamma_2} = \varphi([\gamma_1])\varphi([\gamma_2]).$$

Thus, by Proposition 3.20, $\mathrm{im}(\varphi)$ is a subgroup of $\mathrm{S}(F)$. We define the monodromy group of p associated to the fiber F to be

$$\mathrm{Mon}_F(p) \overset{\text{def}}{=} \mathrm{im}(\varphi).$$

In particular, if p is a finite-sheeted covering map, then $\mathrm{Mon}_F(p)$ is a finite permutation group.

Figure 5.5 shows the covering space which was defined in Section 5.2. We fix a point x on the circle and look at the fiber $F = p^{-1}(x) = \{\ldots, \widetilde{x}_{-2}, \widetilde{x}_{-1}, \widetilde{x}_0, \widetilde{x}_1, \widetilde{x}_2, \ldots\}$. Let γ be a loop based at x such that it runs around the circle only once (red in Figure 5.5). Then we denote unique lifts of γ starting at $\widetilde{x}_i$ by $\widetilde{\gamma}_i$, respectively (blue in Figure 5.5). Then σ_γ just translates all the elements in F by one, i.e.

$$\begin{aligned} \sigma_\gamma\colon F &\to F \\ \widetilde{x}_i &\mapsto \widetilde{x}_{i+1} \end{aligned}$$

The crucial fact is that for different fibers $F_1 = p^{-1}(x_1), F_2 = p^{-1}(x_2)$ the monodromy groups associated to these fibers are isomorphic as permutation groups [15, p. 70]. Thus, in the literature this group is usually denoted $\mathrm{Mon}(p)$ and is called simply the monodromy group of p. In this work we are trying to be rigorous and that's why we keep the notation $\mathrm{Mon}_F(p)$ instead of $\mathrm{Mon}(p)$.

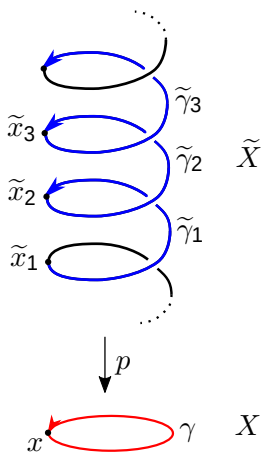


Figure 5.5

Assume now that $\widetilde{X}$ is path-connected. Let's look at the kernel of φ defined in (5.1). These are the homotopy classes of loops based at x which lift to loops starting at $\widetilde{x}$ for every $\widetilde{x}$ 2 F. Fix $\widetilde{x}_1$ 2 F and let $H = p\ (\pi_1(\widetilde{X}, \widetilde{x}_1))$. By Proposition 5.15, H consists exactly of the homotopy classes of loops based at x which lift to loops starting at $\widetilde{x}_1$. Let $\widetilde{x}_2$ be another point in the fiber F. Combining Proposition 5.16 with Proposition 5.15 we conclude that $[\gamma]\ ^1H[\gamma]$ consists exactly of homotopy classes of loops based at x which lift to loops starting at $\widetilde{x}_2$, where $\gamma = p(\widetilde{\gamma})$ for a path $\widetilde{\gamma}$ from $\widetilde{x}_2$ to $\widetilde{x}_1$. Thus we conclude that

$$\ker(\varphi) = \mathrm{Core}_{\pi_1(X,x)}(H) \overset{\text{def}}{=} \bigcap_{[\gamma] 2 \pi_1(X,x)} [\gamma]\ ^1H[\gamma]$$

since to every $\widetilde{x}_2$ 2 F there is a path from $\widetilde{x}_1$ (follows from the fact that $\widetilde{X}$ is path-connected). Thus, by Proposition 3.20,

$$\pi_1(X,x) \Big/ \mathrm{Core}_{\pi_1(X,x)}(H) \overset{\overline{\varphi}}{=} \mathrm{Mon}_F(p) \tag{5.2}$$

Proposition 5.17. *Let $p\colon \widetilde{X}\ !\ X$ be a covering space. Let x 2 X and $F = p\ ^1(x)$. If $\widetilde{X}$ is path-connected, then $\mathrm{Mon}_F(p)$ is transitive.*

Proof. We need to show that for every $\widetilde{x}_1, \widetilde{x}_2$ 2 F there exists σ 2 $\mathrm{Mon}_F(p)$ such that $\sigma(\widetilde{x}_1) = \widetilde{x}_2$. Because $\widetilde{X}$ is path-connected, then there exists a path $\widetilde{\gamma}$ from $\widetilde{x}_1$ to $\widetilde{x}_2$. Let $\gamma = p(\widetilde{\gamma})$ be a loop in X based at x. Thus, for $\sigma = \varphi([\gamma\ ])$ we have $\sigma(\widetilde{x}_1) = \widetilde{\gamma}(1) = \widetilde{x}_2$. □

5.4 Group of deck transformations

Definition 5.18. For a covering space $p\colon \widetilde{X}\ !\ X$ we define $f\colon \widetilde{X}\ !\ \widetilde{X}$ to be a deck transformation of $\widetilde{X}$, if f is a homeomorphism such that $p = pf$.

Notice that such a definition tells us that every deck transformation f acts bijectively on the fiber $p\ ^1(x)$ for every x 2 X.

It is easy to verify that the set of all deck transformations of $\widetilde{X}$ forms a group under the operation of function composition. The group of all deck transformations of a covering space $p\colon \widetilde{X}\ !\ X$ will be denoted as $\mathrm{Deck}(p)$.

By Definition 5.12, f is a lift of p. So, we give the following Proposition.

Proposition 5.19. *Suppose given a covering space $p\colon \widetilde{X}\ !\ X$ with $\widetilde{X}$ path-connected and locally path-connected. Then a lift $\widetilde{p}\colon \widetilde{X}\ !\ \widetilde{X}$ of p, taking $\widetilde{x}_1$ 2 $p\ ^1(x)$ to $\widetilde{x}_2$ 2 $p\ ^1(x)$, exists if and only if $p\ (\pi_1(\widetilde{X}, \widetilde{x}_1))\quad p\ (\pi_1(\widetilde{X}, \widetilde{x}_2))$.*

Proof.) : Since $\widetilde{p}$ is a lift of p, it satisfies $p = p\widetilde{p}$, which implies $p\ = p\ \widetilde{p}$. But then

$$p\ (\pi_1(\widetilde{X}, \widetilde{x}_1)) = p\ (\widetilde{p}\ (\pi_1(\widetilde{X}, \widetilde{x}_1)))\quad p\ (\pi_1(\widetilde{X}, \widetilde{x}_2)),$$

because $\widetilde{p}\ (\pi_1(\widetilde{X}, \widetilde{x}_1))\quad \pi_1(\widetilde{X}, \widetilde{x}_2)$, as $\widetilde{p}$ takes $\widetilde{x}_1$ to $\widetilde{x}_2$.
(: Let $\widetilde{x}$ 2 $\widetilde{X}$ and let γ be a path in $\widetilde{X}$ from $\widetilde{x}_1$ to $\widetilde{x}$. The path $p\gamma$ in X starting at $x_0 = p(\widetilde{x}_1)$ has a unique lift $\widetilde{p\gamma}$ starting at $\widetilde{x}_2$. Define $\widetilde{p}(\widetilde{x}) = \widetilde{p\gamma}(1)$. To show this is

well-defined, independent of the choice of γ, let γ' be another path from $\tilde{x}_1$ to $\tilde{x}$. Then $(p\gamma') \cdot (p\gamma)$ is a loop h in X based at x_0 and by construction, $[h] \in p_*(\pi_1(\tilde{X}, \tilde{x}_1))$. By assumption, $[h] \in p_*(\pi_1(\tilde{X}, \tilde{x}_2))$. This, by Proposition 5.15, means that h lifts to a loop $\tilde{h}$ at $\tilde{x}_2$. By the uniqueness of lifted paths, the first half of $\tilde{h}$ is $\widetilde{p\gamma'}$ and the second half is $\widetilde{p\gamma}$ traversed backwards, with the common midpoint $\widetilde{p\gamma}(1) = \widetilde{p\gamma'}(1)$. This shows that $\tilde{p}$ is well-defined.

To see that $\tilde{p}$ is continuous at every point $\tilde{x}$, let $\tilde{U} \subset \tilde{X}$ be a neighbourhood of $\tilde{p}(\tilde{x})$. Our aim is to prove that there is a neighbourhood V of $\tilde{x}$ such that $\tilde{p}(V) \subset \tilde{U}$. Let $U \subset X$ be a neighbourhood of $p(\tilde{x})$ having a lift $\tilde{U}' \subset \tilde{X}$ containing $\tilde{p}(\tilde{x})$ such that $p\colon \tilde{U}' \to U$ is a homeomorphism (such U and $\tilde{U}'$ exist by the definition of a covering space). Take $\tilde{U}'' = \tilde{U}' \setminus \partial\tilde{U}$ if $\tilde{p}(\tilde{x}) \in \partial\tilde{U}$. Choose a path-connected neighbourhood V of $\tilde{x}$ with $p(V) \subset p(\tilde{U}'')$ (can be done since $p(\tilde{U}'')$ is a neighbourhood of $p(\tilde{x})$ and p is continuous). We now show that $\tilde{p}(V) \subset \tilde{U}''$. For paths from $\tilde{x}_1$ to points $\tilde{x}' \in V$ we can take a fixed path γ from $\tilde{x}_1$ to $\tilde{x}$ followed by paths η in V from $\tilde{x}$ to the points $\tilde{x}'$. The the paths $(p\gamma)\cdot(p\eta)$ in X have lifts $(\widetilde{p\gamma})\cdot(\widetilde{p\eta})$ where $\widetilde{p\eta} = p^{-1}p\eta$ and $p^{-1}\colon p(\tilde{U}'') \to \tilde{U}''$ is the inverse of $p\colon \tilde{U}' \to U$ restricted to $\tilde{U}''$. Thus the end points of lifted paths lie in $\tilde{U}''$, which means that $\tilde{p}(V) \subset \tilde{U}''$. Because $\tilde{U}'' \subset \tilde{U}$, then $\tilde{p}(V) \subset \tilde{U}$. □

The following proposition shows that a lift $\tilde{p}\colon \tilde{X} \to \tilde{X}$ of a covering map is defined by its value at one point.

Proposition 5.20. *Given a covering space $p\colon \tilde{X} \to X$, if $\tilde{X}$ is connected and two lifts $\tilde{p}_1, \tilde{p}_2\colon \tilde{X} \to \tilde{X}$ of p agree at one point of $\tilde{X}$, then $\tilde{p}_1$ and $\tilde{p}_2$ agree on all of $\tilde{X}$.*

Proof. For a point $\tilde{x} \in \tilde{X}$, let U be an open neighbourhood of $p(\tilde{x})$ in X such that $p^{-1}(U)$ is decomposed into disjoint sheets each mapped homeomorphically onto U by p (rewritten definition of a covering space). Let $\tilde{U}_1$ and $\tilde{U}_2$ be the sheets in $p^{-1}(U)$ containing $\tilde{p}_1(\tilde{x})$ and $\tilde{p}_2(\tilde{x})$, respectively. By continuity of $\tilde{p}_1$ and $\tilde{p}_2$ there is an open neighbourhood N of $\tilde{x}$ mapped into $\tilde{U}_1$ by $\tilde{p}_1$ and into $\tilde{U}_2$ by $\tilde{p}_2$. If $\tilde{p}_1(\tilde{x}) \neq \tilde{p}_2(\tilde{x})$ then $\tilde{U}_1 \neq \tilde{U}_2$, hence $\tilde{U}_1$ and $\tilde{U}_2$ are disjoint and $\tilde{p}_1 \neq \tilde{p}_2$ on the whole N. On the other hand, if $\tilde{p}_1(\tilde{x}) = \tilde{p}_2(\tilde{x})$ then $\tilde{U}_1 = \tilde{U}_2$ so $\tilde{p}_1 = \tilde{p}_2$ on the whole N since $p\tilde{p}_1 = p = p\tilde{p}_2$ and p is injective on $\tilde{U}_1 = \tilde{U}_2$.

We now summarize what we've just proved. Let $\tilde{X}_1 \subset \tilde{X}$ be the subset of all $\tilde{x} \in \tilde{X}$ with $\tilde{p}_1(\tilde{x}) \neq \tilde{p}_2(\tilde{x})$. Similarly, let $\tilde{X}_2 \subset \tilde{X}$ be the subset of all $\tilde{x} \in \tilde{X}$ with $\tilde{p}_1(\tilde{x}) = \tilde{p}_2(\tilde{x})$. Easy to see that $\tilde{X}_1$ and $\tilde{X}_2$ are disjoint and that $\tilde{X}_1 \cup \tilde{X}_2 = \tilde{X}$. In the previous paragraph we proved that if $\tilde{X}_1$ contains a point, then it contains an neighbourhood of this point, which is open in $\tilde{X}$. By Proposition 2.33, $\tilde{X}_1$ is open in $\tilde{X}$. Similarly, $\tilde{X}_2$ is open in $\tilde{X}$. So, we can write $\tilde{X}$ as a union of two disjoint open sets. But it's a contradiction since $\tilde{X}$ is connected. That's why $\tilde{X}_1$ is empty and $\tilde{X} = \tilde{X}_2$. □

Proposition 5.21. *Suppose given a covering space $p\colon \tilde{X} \to X$ with $\tilde{X}$ path-connected and locally path-connected. Then a deck transformation $f\colon \tilde{X} \to \tilde{X}$, taking $\tilde{x}_1 \in p^{-1}(x)$ to $\tilde{x}_2 \in p^{-1}(x)$, exists if and only if $p_*(\pi_1(\tilde{X}, \tilde{x}_1)) = p_*(\pi_1(\tilde{X}, \tilde{x}_2))$.*

Proof. If there is a homeomorphism $f\colon \tilde{X} \to \tilde{X}$, taking $\tilde{x}_1$ to $\tilde{x}_2$, then from the two relations $p = pf$ and $p = pf^{-1}$ if follows that $p_*(\pi_1(\tilde{X}, \tilde{x}_1)) = p_*(\pi_1(\tilde{X}, \tilde{x}_2))$. Conversely, suppose that $p_*(\pi_1(\tilde{X}, \tilde{x}_1)) = p_*(\pi_1(\tilde{X}, \tilde{x}_2))$. Using the inclusion $p_*(\pi_1(\tilde{X}, \tilde{x}_1)) \subset p_*(\pi_1(\tilde{X}, \tilde{x}_2))$, by

Proposition 5.19, we can lift p to a map $\widetilde{p}_1\colon \widetilde{X} \to \widetilde{X}$ with $p\widetilde{p}_1 = p$, taking $\widetilde{x}_1$ to $\widetilde{x}_2$. Using the opposite inclusion, we can lift p to a map $\widetilde{p}_2\colon \widetilde{X} \to \widetilde{X}$ with $p\widetilde{p}_2 = p$, taking $\widetilde{x}_2$ to $\widetilde{x}_1$. Consider the maps $\widetilde{p}_2\widetilde{p}_1$ and $\widetilde{p}_1\widetilde{p}_2$. These are lifts of p because $p\widetilde{p}_2\widetilde{p}_1 = p\widetilde{p}_1 = p$ and $p\widetilde{p}_1\widetilde{p}_2 = p\widetilde{p}_2 = p$. Easy to see that $\widetilde{p}_2\widetilde{p}_1$ takes $\widetilde{x}_1$ to $\widetilde{x}_1$ and $\widetilde{p}_1\widetilde{p}_2$ takes $\widetilde{x}_2$ to $\widetilde{x}_2$. According to Proposition 5.20, $\widetilde{p}_2\widetilde{p}_1 = \widetilde{p}_1\widetilde{p}_2 = \mathrm{id}_{\widetilde{X}}$. Thus $\widetilde{p}_1$ and $\widetilde{p}_2$ are mutually inverse maps which means that $\widetilde{p}_1$ is a deck transformation of $\widetilde{X}$ taking $\widetilde{x}_1$ to $\widetilde{x}_2$. □

Remark 5.22. Notice that by Proposition 5.16, the statement $p_*(\pi_1(\widetilde{X}, \widetilde{x}_1)) = p_*(\pi_1(\widetilde{X}, \widetilde{x}_2))$ is equivalent to $[\gamma]\, H[\gamma] = H$, where $H = p_*(\pi_1(\widetilde{X}, \widetilde{x}_1))$ and γ is a loop in X which lifts to a path in $\widetilde{X}$ from $\widetilde{x}_1$ to $\widetilde{x}_2$. By Definition 3.41, this means that $[\gamma]$ lies in the normalizer $\mathrm{N}_{\pi_1(X,x)}(H)$ of H.

The following proposition reveals the structure of $\mathrm{Deck}(p)$. Unfortunately, A. Hatcher gives just a sketched proof of this proposition. We, however, tried to make the proof as detailed as possible.

Proposition 5.23. *Let $p\colon \widetilde{X} \to X$ be a covering space with $\widetilde{X}$ path-connected and locally path-connected. Fix $\widetilde{x}_1 \in \widetilde{X}$ and let H be the subgroup $p_*(\pi_1(\widetilde{X}, \widetilde{x}_1)) \subseteq \pi_1(X, x)$. Then*

$$\mathrm{Deck}(p) = \mathrm{N}_{\pi_1(X,x)}(H)/H.$$

Proof. Define a map

$$\varphi\colon \mathrm{N}_{\pi_1(X,x)}(H) \to \mathrm{Deck}(p)$$
$$[\gamma] \mapsto \left(f\colon \widetilde{x} \mapsto \alpha \cdot \widetilde{\gamma} \cdot \alpha\,(1)\right)$$

where $\alpha = p(\widetilde{\alpha}_1)$ for some path $\widetilde{\alpha}_1$ in $\widetilde{X}$ from $\widetilde{x}_1$ to $\widetilde{x}$ and $\alpha \cdot \widetilde{\gamma} \cdot \alpha$ is a unique lift of $\alpha \cdot \gamma \cdot \alpha$ starting at $\widetilde{x}$.

We prove that φ is well-defined. 1) At first, we show that it doesn't depend on the choice of class representative γ. For this fix $\widetilde{x}$ and $\widetilde{\alpha}_1$ and denote $\alpha = p(\widetilde{\alpha}_1)$. Let $[\gamma_1] = [\gamma_2] \in \mathrm{N}_{\pi_1(X,x)}(H)$. It means there is a homotopy f_t between γ_1 and γ_2, fixing the endpoints. Let $\alpha \cdot \widetilde{\gamma}_1 \cdot \alpha$ and $\alpha \cdot \widetilde{\gamma}_2 \cdot \alpha$ be unique lifts of $\alpha \cdot \gamma_1 \cdot \alpha$ and $\alpha \cdot \gamma_2 \cdot \alpha$, respectively, starting at $\widetilde{x}$. Easy to see that $\alpha \cdot \gamma_1 \cdot \alpha$ and $\alpha \cdot \gamma_2 \cdot \alpha$ are homotopic r.t.e. (just take a homotopy which fixes α and α pointwise and transforms γ_1 to γ_2 in the same way as f_t). By Remark 5.14, $\alpha \cdot \widetilde{\gamma}_1 \cdot \alpha$ and $\alpha \cdot \widetilde{\gamma}_2 \cdot \alpha$ are homotopic r.t.e, which

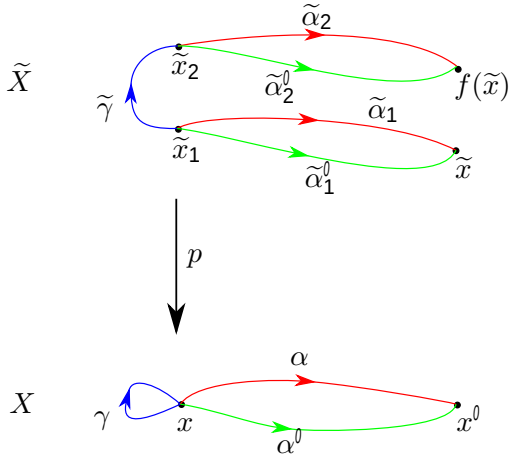


Figure 5.6

means there endpoints are the same. 2) Now we prove that φ doesn't depend on the choice of a path $\widetilde{\alpha}_1$ (see Figure 5.6). For this, fix γ and $\widetilde{x}$ and let $\widetilde{\alpha}_1$, $\widetilde{\alpha}_1'$ be two different paths in $\widetilde{X}$ from $\widetilde{x}_1$ to $\widetilde{x}$. Let also $\alpha = p(\widetilde{\alpha}_1)$, $\alpha' = p(\widetilde{\alpha}_1')$. We need to prove that $\alpha \cdot \overline{\gamma} \cdot \alpha\ (1) = \alpha' \cdot \overline{\gamma} \cdot \alpha'\ (1)$. Let $\widetilde{\gamma}$ be a unique lift of γ starting at $\widetilde{x}_1$. Denote $\widetilde{x}_2 = \widetilde{\gamma}(1)$. Let $\widetilde{\alpha}_2$ and $\widetilde{\alpha}_2'$ be lifts of α and α', respectively, starting at $\widetilde{x}_2$. It is easy to see that $\widetilde{\alpha}_1 \cdot \widetilde{\gamma} \cdot \widetilde{\alpha}_2$ and $\widetilde{\alpha}_1' \cdot \widetilde{\gamma} \cdot \widetilde{\alpha}_2'$ are unique lifts of $\alpha \cdot \gamma \cdot \alpha$ and $\alpha' \cdot \gamma \cdot \alpha'$, respectively. We now prove that these two lifts end at the same point. For this it is enough to prove that $\widetilde{\alpha}_2$ and $\widetilde{\alpha}_2'$ end at the same point. Since $\gamma \in \mathrm{N}_{\pi_1(X,x)}(H)$, then, according to Remark 5.22, $p_*(\pi_1(\widetilde{X}, \widetilde{x}_1)) = p_*(\pi_1(\widetilde{X}, \widetilde{x}_2))$, which by Proposition 5.15, means that the loop $\alpha \cdot \alpha'$ lifts to a loop based at $\widetilde{x}_2$. By homotopy lifting property, this loop has the form $\widetilde{\alpha}_2 \cdot \widetilde{\alpha}_2'$, and then $\widetilde{\alpha}_2$ and $\widetilde{\alpha}_2'$ end at the same point. 3) We show that φ indeed produces a deck transformation of the covering space $p\colon \widetilde{X} \to X$. For this just notice that $\widetilde{x} \mapsto \alpha \cdot \overline{\gamma}^{-1} \cdot \alpha\ (1)$ is exactly the map constructed in the proof of Proposition 5.19, because in the point 2) we have shown that $\alpha \cdot \overline{\gamma}^{-1} \cdot \alpha = \widetilde{\alpha}_1 \cdot \widetilde{\gamma} \cdot \widetilde{\alpha}_2$. And since $p_*(\pi_1(\widetilde{X}, \widetilde{x}_1)) = p_*(\pi_1(\widetilde{X}, \widetilde{x}_2))$, then, by Proposition 5.21, the map $\varphi([\gamma])$ is indeed a deck transformation. It is also easy to notice that $\varphi([\gamma])$ sends $\widetilde{x}_1$ to $\widetilde{x}_2$. We claim now that φ is a group homomorphism (see Figure 5.7). Let $[\gamma_1^{-1}], [\gamma_2^{-1}] \in \mathrm{N}_{\pi_1(X,x)}(H)$. Then

$$\varphi([\gamma_1][\gamma_2]) = \varphi([\gamma_1 \cdot \gamma_2]) = \varphi([(\gamma_2 \cdot \gamma_1)]) = \left(\widetilde{x} \mapsto \alpha \cdot \overline{\gamma_2 \cdot \gamma_1} \cdot \alpha\ (1)\right) =$$

$$= \left(\widetilde{x} \mapsto \alpha \cdot \gamma_2 \cdot \overline{\alpha \cdot \alpha} \cdot \gamma_1 \cdot \alpha\ (1)\right) = \left(\widetilde{x} \mapsto \alpha \cdot \overline{\gamma_2} \cdot \alpha \cdot \alpha \cdot \overline{\gamma_1} \cdot \alpha\ (1)\right),$$

where $\alpha \cdot \overline{\gamma_2} \cdot \alpha$ and $\alpha \cdot \overline{\gamma_1} \cdot \alpha$ are unique lifts of $\alpha \cdot \gamma_2 \cdot \alpha$ and $\alpha \cdot \gamma_1 \cdot \alpha$ starting at $\widetilde{x}$ and $\alpha \cdot \overline{\gamma_2} \cdot \alpha\ (1)$, respectively. But now,

$$\left(\widetilde{x} \mapsto \alpha \cdot \overline{\gamma_2} \cdot \alpha \cdot \alpha \cdot \overline{\gamma_1} \cdot \alpha\ (1)\right) = \left(\widetilde{x} \mapsto \alpha \cdot \overline{\gamma_1} \cdot \alpha\ (1)\right)\left(\widetilde{x} \mapsto \alpha \cdot \overline{\gamma_2} \cdot \alpha\ (1)\right) =$$

$$= \varphi([\gamma_1])\varphi([\gamma_2]).$$

We prove that $\ker(\varphi) = H$. The loops γ^{-1} with $[\gamma] \in H$ are exactly the loops in $\pi_1(X, x)$ which lift to loops in $\widetilde{X}$ based at $\widetilde{x}_1$. That's why $\varphi([\gamma]) = \mathrm{id}_{\widetilde{X}}$. Conversely, take $[\gamma] \in \mathrm{N}_{\pi_1(X,x)}(H)$ with $[\gamma] \notin H$. Then γ lifts to a path starting at $\widetilde{x}_1$, which is not a loop at $\widetilde{x}_1$. Then $\varphi([\gamma])$ sends $\widetilde{x}$ to $f(\widetilde{x}) \neq \widetilde{x}$, because if $f(\widetilde{x}) = \widetilde{x}$ held true, then there would exist two different lifts $\widetilde{\alpha}_2$, $\widetilde{\alpha}_1$ of α starting at the same point $\widetilde{x}$, which would contradict to the homotopy lifting property. Hence $\varphi([\gamma]) \neq \mathrm{id}_{\widetilde{X}}$.

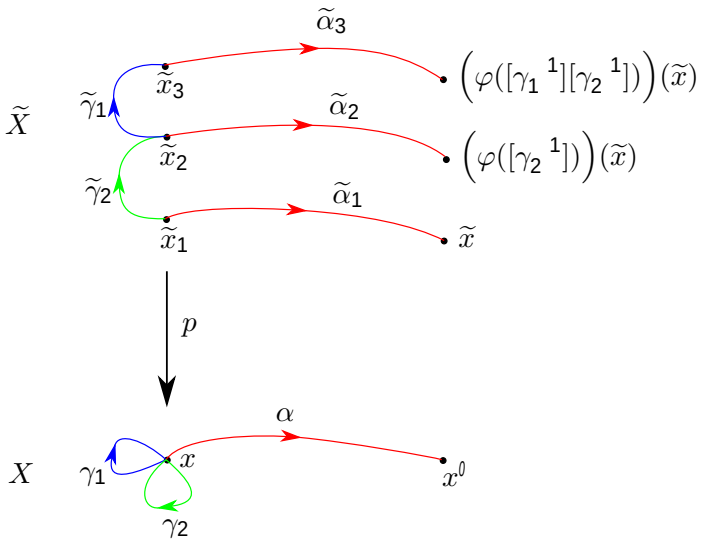


Figure 5.7

It remains to prove that φ is surjective. Take any $f \in \mathrm{Deck}(p)$. Let $\widetilde{\gamma}$ be a path in $\widetilde{X}$ from $\widetilde{x}_1$ to $f(\widetilde{x}_1)$. By Proposition 5.21, there holds

$$p\ (\pi_1(\widetilde{X}, \widetilde{x}_1)) = p\ (\pi_1(\widetilde{X}, f(\widetilde{x}_1))),$$

which, by Remark 5.22, means that $\gamma = p\widetilde{\gamma}$ represents the class in the normalizer $\mathrm{N}_{\pi_1(X,x)}(H)$. Hence, $\varphi([\gamma\ ])$ is a deck transformation sending $\widetilde{x}_1$ to $f(\widetilde{x}_1)$. But then, by Proposition 5.20, $\varphi([\gamma\ ])$ coincides with f.

We now have a group homomorphism φ which is surjective and has $\ker(\varphi) = H$. By Proposition 3.20, it follows that

$$\mathrm{N}_{\pi_1(X,x)}(H)/H \overset{\overline{\varphi}}{=} \mathrm{Deck}(p) \tag{5.3}$$

We can now fix $x \in X$ and look at how $\mathrm{Deck}(p)$ acts on the fiber $F = p^{-1}(x)$. Because f is bijective on the whole $\widetilde{X}$, it acts bijectively on F. We define a group homomorphism

$$\begin{aligned} \psi\colon \mathrm{Deck}(p) &\to \mathrm{S}(F) \\ f &\mapsto f\big|_F \end{aligned}$$

It is easy to see that ψ is injective because, by Proposition 5.20, if two deck transformations coincide on the fiber, they coincide on the whole $\widetilde{X}$. The action of $\mathrm{Deck}(p)$ on the fiber F will be denoted as $\mathrm{Deck}_F(p) \overset{\text{def}}{=} \psi(\mathrm{Deck}(p))$. And because ψ is injective, then

$$\mathrm{Deck}(p) \overset{\psi}{=} \mathrm{Deck}_F(p) \tag{5.4}$$

Since $\mathrm{S}(F)$ is finite, then $\mathrm{Deck}(p)$ is finite.

5.5 Relations between the monodromy group and the group of deck transformations

It is known [7, Proposition 1.4] that

$$\mathrm{Deck}_F(p) = \mathrm{C}_{\mathrm{S}(F)}\Big(\mathrm{Mon}_F(p)\Big) \tag{5.5}$$

where $\mathrm{C}_G(H)$ denotes the centralizer of H in G. Our aim here is to give a detailed proof of (5.5) which doesn't require a knowledge of the fundamental theorem of covering spaces. The proof is based on basic group theory (see Section 3.4).

The monodromy $\mathrm{Mon}_F(p) \subset \mathrm{S}(F)$ is a transitive permutation group (associate with G from Section 3.4), the stabilizer $\mathrm{Stab}_{\mathrm{Mon}_F(p)}(\widetilde{x})$ of some element $\widetilde{x} \in F$ (associate with G_x) and the normalizer $\mathrm{N}_{\mathrm{Mon}_F(p)}\Big(\mathrm{Stab}_{\mathrm{Mon}_F(p)}(\widetilde{x})\Big)$ (associate with NG_x). For simplicity, we denote $\mathrm{Stab}_{\mathrm{Mon}_F(p)}(\widetilde{x})$ by $M_{\widetilde{x}}$. We can create exactly the same diagram as in (3.8):

$$\begin{array}{ccccc} \mathrm{Mon}_F(p) & \xrightarrow{\rho_L} & \mathrm{S}\Big(\mathrm{Mon}_F(p)\Big/M_{\widetilde{x}}\Big) & \xleftarrow{\overline{\rho_R}} & \mathrm{N}_{\mathrm{Mon}_F(p)}\Big(M_{\widetilde{x}}\Big)\Big/M_{\widetilde{x}} \\ & \searrow^{\mathrm{id}_M} & \Big\downarrow{=}\ \Big\downarrow\phi & & \\ & & \mathrm{S}(F) & & \end{array} \tag{5.6}$$

where id_M is the identity map on $\mathsf{Mon}_F(p)$. By Proposition 3.47, we have

$$\mathrm{C}_{\mathsf{S}\left(\mathsf{Mon}_F(p)\Big/M_{\widetilde{x}}\right)}\left(\rho_L\big(\mathsf{Mon}_F(p)\big)\right) = \overline{\rho_R}\left(\mathrm{N}_{\mathsf{Mon}_F(p)}\left(M_{\widetilde{x}}\right)\Big/M_{\widetilde{x}}\right) \tag{5.7}$$

We would like to add the group $\mathsf{Deck}_F(p)$ somewhere to diagram 5.6 without violating its commutativity. We propose the following diagram:

$$\begin{array}{ccccc}
\mathsf{Mon}_F(p) & \xhookrightarrow{\rho_L} & \mathsf{S}\left(\mathsf{Mon}_F(p)\Big/M_{\widetilde{x}}\right) & \xhookleftarrow{\overline{\rho_R}} & \mathrm{N}_{\mathsf{Mon}_F(p)}\left(M_{\widetilde{x}}\right)\Big/M_{\widetilde{x}} \\
\Big\| & & \Big\| & & \Big\| \overline{\beta} \\
\mathsf{Mon}_F(p) & \xhookrightarrow{\mathsf{id}_M} & \mathsf{S}(F) & \xhookleftarrow{\mathsf{id}_D} & \mathrm{N}_{\pi_1(X,x)}(H)/H \\
 & & & & \Big\| \psi\overline{\varphi} \\
 & & & & \mathsf{Deck}_F(p)
\end{array} \tag{5.8}$$

where $H = p_*(\pi_1(\widetilde{X},\widetilde{x}))$, $\overline{\varphi}$ and ψ are the isomorphisms from (5.3) and (5.4), respectively, and id_D is the identity map on $\mathsf{Deck}_F(p)$. If it was possible to do that (i.e. find such an isomorphism $\overline{\beta}$ and show that this diagram is commutative), the intuition would tell us that

$$\mathrm{C}_{\mathsf{S}(F)}\left(\mathsf{Mon}_F(p)\right) = \mathsf{Deck}_F(p)$$

holds true. Luckily, such an isomorphism $\overline{\beta}$, which makes the diagram (5.8) commutative, exists. We prove the following proposition which will help us to find $\overline{\beta}$.

Proposition 5.24. *Let G be a group and $H^0 \quad H \quad G$ be a chain of subgroups with H^0 normal in G. Let $\widetilde{G} = G/H^0$ be the quotient group and $\widetilde{H} = H/H^0$ be a subgroup of $\widetilde{G}$. Then:*

1. $\mathrm{N}_{\widetilde{G}}(\widetilde{H}) = \mathrm{N}_G(H)/H^0$.
2. $\mathrm{N}_{\widetilde{G}}(\widetilde{H})/\widetilde{H} \overset{\overline{\beta}}{=} \mathrm{N}_G(H)/H$.

Proof. 1. By definition,

$$\mathrm{N}_G(H) = \{g \in G \mid ghg^{-1} \in H \ \ \forall h \in H\},$$

$$\mathrm{N}_{\widetilde{G}}(\widetilde{H}) = \{gH^0 \in \widetilde{G} \mid gH^0hH^0g^{-1}H^0 \in \widetilde{H} \ \ \forall h \in H\},$$

$$\mathrm{N}_G(H)/H^0 = \{gH^0 \in \widetilde{G} \mid ghg^{-1} \in H \ \ \forall h \in H\}.$$

Because H^0 is normal in G, then

$$gH^0hH^0g^{-1}H^0 = ghH^0g^{-1}H^0 = ghg^{-1}H^0.$$

The condition $gH^0hH^0g^{-1}H^0 \in \widetilde{H}\ \forall h \in H$ defining $\mathrm{N}_{\widetilde{G}}(\widetilde{H})$ is then equivalent to $ghg^{-1}H^0 \in \widetilde{H}\ \forall h \in H$, or just $ghg^{-1} \in H\ \forall h \in H$. We see that this coincides with the condition

defining $\mathrm{N}_G(H)/H^0$.
2. By Proposition 3.24, we have

$$\mathrm{N}_{\widetilde{G}}(\widetilde{H})/\widetilde{H} = (\mathrm{N}_G(H)/H^0)/(H/H^0) \overset{\overline{\beta}}{=} \mathrm{N}_G(H)/H.$$

□

Using the notation of Proposition 5.24, we can associate $\pi_1(X,x)$ with G, $H = p\ (\pi_1(\widetilde{X},\widetilde{x}))$ with H and $\mathrm{Core}_{\pi_1(X,x)}(H)$ with H^0. To apply an isomorphism $\overline{\beta}$ from Proposition 5.24 to diagram (5.8) we only need to prove that

$$M_{\widetilde{x}} = H \Big/ \mathrm{Core}_{\pi_1(X,x)}(H).$$

But it is straightforward because, by Proposition 5.15, $M_{\widetilde{x}}$ consists of classes of homotopy classes of loops at x which lift to loops at $\widetilde{x}$.

We now show that diagram (5.8) is commutative.

Proposition 5.25. *Diagram* (5.8) *is commutative.*

Proof. It was proven earlier that the left part of the diagram commutes, i.e.

$$\phi\rho_L = \iota_M \mathrm{id}.$$

It remains to show that

$$\phi\overline{\rho_R} = \iota_D\psi\overline{\varphi}\overline{\beta}.$$

Take $nM_{\widetilde{x}}$ 2 $\mathrm{N}_{\mathtt{Mon}_F(p)}\Big(M_{\widetilde{x}}\Big)\Big/M_{\widetilde{x}}$. Then

$$\overline{\rho_R}(nM_{\widetilde{x}}) = \Big(m^0M_{\widetilde{x}}\ \mathbin{7!}\ m^0n\ {}^1M_{\widetilde{x}}\Big).$$

Applying ϕ to it we get

$$\phi(\overline{\rho_R}(nM_{\widetilde{x}})) = \Big(\widetilde{x}^0\ \mathbin{7!}\ (\eta\overline{\rho_R}(nM_{\widetilde{x}})\eta\ {}^1)(\widetilde{x}^0)\Big) = \Big(\widetilde{x}^0\ \mathbin{7!}\ mn\ {}^1(\widetilde{x})\Big),$$

where $m(\widetilde{x}) = \widetilde{x}^0$, or $\widetilde{x} = m\ {}^1(\widetilde{x}^0)$. So the latter map can be rewritten as

$$\phi(\overline{\rho_R}(nM_{\widetilde{x}})) = \Big(\widetilde{x}^0\ \mathbin{7!}\ mn\ {}^1m\ {}^1(\widetilde{x}^0)\Big), \quad m\ {}^1(\widetilde{x}^0) = \widetilde{x} \tag{5.9}$$

Let's now find out how $\iota_D\psi\overline{\varphi}\overline{\beta}$ acts on $nM_{\widetilde{x}}$. Remember that n is an element in $\mathrm{N}_{\mathtt{Mon}_F(p)}\Big(M_{\widetilde{x}}\Big)$, which means, by Proposition 5.24, it can be written as $\big[\gamma\big]\mathrm{Core}_{\pi_1(X,x)}(H)$ for some homotopy class of loops $\big[\gamma\big]$ 2 $\mathrm{N}_{\pi_1(X,x)}(H)$. Then,

$$nM_{\widetilde{x}} = \Big([\gamma]\mathrm{Core}_{\pi_1(X,x)}(H)\Big)M_{\widetilde{x}} \overset{\overline{\beta}}{\longmapsto} \big[\gamma\big]H \overset{\iota_D\psi\overline{\varphi}}{\longmapsto} \Big(\widetilde{x}^0\ \mathbin{7!}\ \alpha\ \ \dot{\overline{\gamma}}\ \ \cdot\alpha\ (1)\Big)$$

where $\alpha = p(\widetilde{\alpha})$ for some path $\widetilde{\alpha}$ from $\widetilde{x}^0$ to $\widetilde{x}$ and with the lift $\alpha\ \ \dot{\overline{\gamma}}\ \ \cdot\alpha$ starting at $\widetilde{x}^0$. But this is exactly what does the map from (5.9). □

Before proving the main result of this chapter we will need one more result concerning how the centralizer of a subgroup changes under a group isomorphism.

Proposition 5.26. *Let* $\phi\colon G \xrightarrow{\sim} G'$ *be a group isomorphism and* $H \leq G$ *be a subgroup. Then*

$$\phi\Big(\mathrm{C}_G(H)\Big) = \mathrm{C}_{G'}\Big(\phi(H)\Big).$$

Proof. $\subseteq$: Take $g' \in \phi\Big(\mathrm{C}_G(H)\Big)$. Then $g' = \phi(g)$ with

$$gh = hg \quad \forall h \in H.$$

Applying ϕ to both sides of the above equality we get

$$g'\phi(h) = \phi(g)\phi(h) = \phi(gh) = \phi(hg) = \phi(h)\phi(g) = \phi(h)g' \quad \forall h \in H,$$

which means that $g' \in \mathrm{C}_{G'}\Big(\phi(H)\Big)$.

$\supseteq$: Take $g' \in \mathrm{C}_{G'}\Big(\phi(H)\Big)$. Then

$$g'\phi(h) = \phi(h)g', \quad \forall h \in H.$$

Applying an inverse isomorphism ϕ^{-1} to both sides we get

$$\phi^{-1}(g')h = \phi^{-1}(g'\phi(h)) = \phi^{-1}(\phi(h)g') = h\phi^{-1}(g'), \quad \forall h \in H.$$

This means that

$$\phi^{-1}(g') \in \mathrm{C}_G(H),$$

or that

$$g' \in \phi\Big(\mathrm{C}_G(H)\Big).$$

□

We are now ready to prove the main result of this chapter.

Proposition 5.27. *In diagram* (5.8)*,*

$$\mathrm{C}_{\mathrm{S}(F)}\Big(\mathtt{Mon}_F(p)\Big) = \mathtt{Deck}_F(p)$$

holds true.

Proof. Applying ϕ to both sides of Equation (5.7) we obtain the desired result using Proposition 5.26 and the fact that diagram (5.8) is commutative. □

6 Classical Galois Theory

As we know from Chapter 2, Definition 2.20, a field is a ring where every nonzero element has a multiplicative inverse. By Definition 4.13, we know that given an irreducible a□ne variety X C^n, we can construct a field C(X) which is called the function field of X. By Proposition 4.42, a dominant rational map φ: X 99K Y between irreducible a□ne varieties induces an inclusion φ : C(Y) ‹! C(X) of function fields, so that φ (C(Y)) lies inside C(X). A pair of fields φ (C(Y)) C(X) is called a field extension. At this point we could move to field theory to reveal some interesting properties of φ. It turns out that combining field theory with group theory makes our investigations easier. This is exactly what classical Galois theory (discovered by Évariste Galois in 19th century) does: it reformulates certain problems in field theory in the language of group theory, which makes these problems easier to solve. In Chapters 7 and 8 we will show how classical Galois theory can be used to reveal useful properties of dominant rational maps.

6.1 Field extensions

Every field F has only two ideals: $f0_F g$ and F itself. Indeed, if there is a nontrivial proper ideal I F, then for some nonzero a 2 I we have a a $^1 = 1_F$ 2 I and thus every element from F lies in I. This implies that every ring homomorphism of fields φ: F ! L, by Proposition 2.14, has trivial kernel, i.e. $F = \varphi(F)$. Hence L contains a field isomorphic F.

Example 6.1. The field C of complex numbers can be defined as the quotient ring

$$\mathsf{C} \stackrel{\text{def}}{=} \mathsf{R}[x]/hx^2 + 1i.$$

Then we can embed R into C via the following ring homomorphism:

$$\begin{aligned} \varphi\colon \mathsf{R} &\text{ ‹! } \mathsf{C} \\ a &\text{ 7! } a + hx^2 + 1i \end{aligned}$$

Then φ(R) is isomorphic to R.

Definition 6.2. Given a ring homomorphism of fields φ: F ‹! L, we will usually identify F with its image $\varphi(F)$ and write F L. In other words, we redefine the subset expression F L for fields F, L in the following way:

$$F \quad L \quad ( ) \quad \text{there is a ring homomorphism } \varphi\colon F \text{ ‹! } L$$

and we say that F L is a field extension. We will also say that F is a subfield of L.

Given a field extension $F \subseteq L$, elements of the larger field L can relate to the smaller field F in two different ways.

Definition 6.3. Let $F \subseteq L$ be a field extension, and let $\alpha \in L$. Then α is **algebraic** over F if there is a nonconstant polynomial $f \in F[x]$ such that $f(\alpha) = 0$. If α is not algebraic over F, then α is **transcendental** over F.

Example 6.4. The number $\sqrt{2} \in \mathbb{R}$ is algebraic over $\mathbb{Q}$, since $\sqrt{2}$ is a root of $x^2 - 2 \in \mathbb{Q}[x]$. Also, $i \in \mathbb{C}$ is algebraic over $\mathbb{Q}$, since it is a root of $x^2 + 1 \in \mathbb{Q}[x]$.

When $\alpha \in L$ is algebraic over F, there may be many nonconstant polynomials in $F[x]$ with α as a root. One of these polynomials is especially nice.

Proposition 6.5. *If $\alpha \in L$ is algebraic over F, then there is a unique nonconstant monic polynomial $p \in F[x]$ with the following two properties:*

(a) α is a root of p, i.e., $p(\alpha) = 0$

(b) If $f \in F[x]$ is any polynomial with α as a root, then $f = q \cdot p$ for some $q \in F[x]$.

Proof. Among all nonconstant polynomials in $F[x]$ with α as a root, there must be one of smallest degree. Pick one such polynomial and call it p. Multiplying by a constant if necessary, we may assume that p is monic.

This polynomial certainly satisfies (a). As for (b), suppose that $f(\alpha) = 0$ for some $f \in F[x]$. The division algorithm gives us polynomials $q, r \in F[x]$ such that

$$f = q \cdot p + r, \quad r = 0 \text{ or } \deg(r) < \deg(p).$$

Evaluating the above equation at α gives

$$0 = f(\alpha) = q(\alpha) \cdot p(\alpha) + r(\alpha) = q(\alpha) \cdot 0 + r(\alpha) = r(\alpha).$$

If r had strictly smaller degree than p, this would contradict the definition of p. Hence $r = 0$ and (b) follows.

Finally, to prove uniqueness of p, suppose that another monic polynomial $\tilde{p}$ satisfies properties (a) and (b). Then applying (b) for both p and $\tilde{p}$ we get that $\tilde{p}$ divides p and p divides $\tilde{p}$. Because they are monic, it means they should be equal. □

Definition 6.6. Let $\alpha \in L$. If α is algebraic over F, then the polynomial p in Proposition 6.5 is called the **minimal polynomial** of α over F.

There is also another way to think about the minimal polynomial.

Proposition 6.7. *Let $\alpha \in L$ be algebraic over F, and let $p \in F[x]$ be its minimal polynomial. If $f \in F[x]$ is a nonconstant monic polynomial with $f(\alpha) = 0$, then*

$$f = p \iff f \text{ is irreducible over } F.$$

Proof.) : If $p = gh$, where some $g, h \in F[x]$ have strictly smaller degree than p, then $0 = p(\alpha) = g(\alpha)h(\alpha)$ would imply $g(\alpha) = 0$ or $h(\alpha) = 0$. This, however, contradicts the minimality of degree of p.
(: If f is irreducible, then accordint to the proof of Proposition 6.5, p divides f, so that $f = ph$ for some $h \in F[x]$. Since f is irreducible and p is nonconstant, h must be constant. Because f is monic, then $f = p$. □

Example 6.8. Since $\sqrt{2}$ is irrational, the minimal polynomial of $\sqrt{2}$ over Q has degree at least two. It is easy to see that $x^2 - 2$ has $\sqrt{2}$ as a root, and so is the minimal polynomial of $\sqrt{2}$ over Q.

We next show that, given a field extension $F \subseteq L$ and elements $\alpha_1, \ldots, \alpha_n \in L$, we can create a subfield of L which contains F and $\alpha_1, \ldots, \alpha_n$. We define

$$F[\alpha_1, \ldots, \alpha_n] = \Big\{ h(\alpha_1, \ldots, \alpha_n) \mid h \in F[x_1, \ldots, x_n] \Big\},$$

where $F[x_1, \ldots, x_n]$ is the ring of polynomials over F in n variables. Hence $F[\alpha_1, \ldots, \alpha_n]$ consists of all polynomial expressions in L that can be formed using $\alpha_1, \ldots, \alpha_n$ with coefficients in F. Let

$$F(\alpha_1, \ldots, \alpha_n) = \left\{ \frac{\alpha}{\beta} \;\middle|\; \alpha, \beta \in F[\alpha_1, \ldots, \alpha_n], \beta \neq 0 \right\},$$

Thus, $F(\alpha_1, \ldots, \alpha_n)$ is the set of all rational expressions in α_i with coefficients in F. We say that $F(\alpha_1, \ldots, \alpha_n)$ is obtained from F by adjoining $\alpha_1, \ldots, \alpha_n \in L$. We can characterize $F(\alpha_1, \ldots, \alpha_n)$ as the smallest subfield of L containing F and $\alpha_1, \ldots, \alpha_n$, meaning, if K is another subfield of L which contains F and $\alpha_1, \ldots, \alpha_n$, then $F(\alpha_1, \ldots, \alpha_n) \subseteq K$. This is because K is closed under addition, multiplication and taking the inverse of a nonzero element.

Remark 6.9. Given a field extension $F \subseteq L$ and an algebraic element $\alpha \in L$ over F, we can construct a map

$$\begin{aligned} \varphi \colon F[x] &\to F[\alpha] \\ f(x) &\mapsto f(\alpha) \end{aligned}$$

It is easy to verify that φ is a ring homomorphism. The kernel of this map is the set of all polynomials which has α as a root. The proof of Proposition 6.5 shows that every such polynomial is a multiple of the minimal polynomial p of α over F. This map is surjective because every element $F[\alpha]$ is a polynomial in α. Hence, by Proposition 2.14,

$$F[x]/\langle p \rangle = F[\alpha].$$

Because p is irreducible (see Proposition 6.7), then $\langle p \rangle$ is a maximal ideal. So, $F[x]/\langle p \rangle$ is a field and then so is $F[\alpha]$. Since $F(\alpha)$ is the smallest field containing F and α, we must have $F[\alpha] = F(\alpha)$.

When F is a subfield of a field L, there is one bit of structure that hasn't been used yet. We know that L is an Abelian group under addition (see Definition 2.1). We can also multiply every element of L by any element of F. It is easy to check that these two operations give L a structure of a vector space over F.

Definition 6.10. Let $F \quad L$ be a field extension. Then L is a finite extension of F if L is a finite-dimensional vector space over F. The degree of L over F, denoted $[L : F]$, is defined as follows:

$$[L : F] = \begin{cases} \dim_F L, & \text{if } L \text{ is a finite extension of } F, \\ 1, & \text{otherwise,} \end{cases}$$

where $\dim_F L$ is the dimension of L as a vector space over F.

Example 6.11. Recall Example 6.1. Every element of C (a coset $f(x) + hx^2 + 1i$ for $f(x)$ 2 R$[x]$) can be written uniquely as $ax + b + hx^2 + 1i$ for a, b 2 R. This is because the remainder of $f(x)$ after division by $x^2 + 1$ is unique and has degree less than 2. So, every element in C can be written as a linear combination of $1 + hx^2 + 1i$ and $x + hx^2 + 1i$ with coe□cients in R, which makes the set $f1+hx^2+1i, x+hx^2+1ig$ a basis of C over R. Further, we will denote these two elements as 1 and i, respectively.

Proposition 6.12. *Suppose that $F \quad L$ is a field extension and α 2 L.*

(a) $[F(\alpha) : F] < 1$ if and only if α is algebraic over F.

(b) Let α be algebraic over F. If n is the degree of the minimal polynomial of α over F, then $1, \alpha, \ldots, \alpha^{n \ 1}$ form a basis of $F(\alpha)$ over F. Thus $[F(\alpha) : F] = n$.

Proof. Let $n = [F(\alpha) : F]$ be the dimension of $F(\alpha)$ over F. Then any collection of $n + 1$ elements of $F(\alpha)$ is linearly dependent over F. In particular, $1, \alpha, \ldots, \alpha^n$ are linearly dependent over F. This means there are $a_0, \ldots, a_n$ 2 F, not all zero, such that

$$a_0 + a_1\alpha + \quad + a_n\alpha^n = 0.$$

It follows that α is a root of

$$a_0 + a_1x + \quad + a_nx^n \text{ 2 } F[x],$$

which is nonzero, since the a_i's are not all zero. Hence α is algebraic over F.
Conversely, let α be algebraic over F with minimal polynomial p, where $n = \deg(p)$. By Remark 6.9, $F[\alpha] = F(\alpha)$, so every element of $F(\alpha)$ is of the form $g(\alpha)$ for some g 2 $F[x]$. Dividing g by p gives

$$g = qp + a_0 + a_1x + \quad + a_{n \ 1}x^{n \ 1},$$

where q 2 $F[x]$ and $a_0, \ldots, a_{n \ 1}$ 2 F, and evaluating this at $x = \alpha$ yields

$$g(\alpha) = a_0 + a_1\alpha + \quad + a_{n \ 1}\alpha^{n \ 1},$$

since $p(\alpha) = 0$. Thus $1, \alpha, \ldots, \alpha^{n \ 1}$ span $F(\alpha)$ over F. To show linear independence, suppose that

$$0 = b_0 + b_1\alpha + \quad + b_{n \ 1}\alpha^{n \ 1}$$

for some $b_0, b_1, \ldots, b_{n \ 1}$ 2 F. Then α is a root of $b_0 + b_1x + \quad + b_{n \ 1}x^{n \ 1}$ 2 $F[x]$. Since the minimal polynomial p has degree n, this must be the zero polynomial. Hence $b_i = 0$ for all i, and linear independence is proved. Then $[F(\alpha) : F] = n$ follows from Definition 6.10. □

Proposition 6.13. *Suppose that we have fields* $F \quad K \quad L$. *If* $[L : F] < 1$, *then* $[L : K] < 1$ *and* $[K : F] < 1$. *Moreover, if* $[L : F] < 1$, *then* $[L : F] = [L : K][K : F]$.

Proof. Since L is a finite-dimensional vector space over F, we can pick a basis $\gamma_1, \ldots, \gamma_N$ of L over F. Then:

1. One easily sees that $K \quad L$ is a subspace of L over F. Since L has finite dimension over F, so does any subspace. Hence $[K : F] < 1$.
2. Take $\alpha \; 2 \; L$. Since $\gamma_1, \ldots, \gamma_N$ span L over F, $\alpha = \sum_{i=1}^{N} a_i \gamma_i$, where $a_i \; 2 \; F$. Since $F \quad K$, we can consider this as a linear combination with coe□cients in K. Thus L is spanned over K by a finite set, so that $[L : K] < 1$.

For the moreover part, just pick bases $\alpha_1, \ldots, \alpha_m$ of K over F and $\beta_1, \ldots, \beta_n$ of L over K. It is straightforward to prove that the mn products

$$\alpha_i \beta_j, \quad 1 \quad i \quad m, 1 \quad j \quad n,$$

form a basis of L over F. □

Definition 6.14. A field extension $F \quad L$ is algebraic if every element of L is algebraic over F.

The next proposition shows that every finite extension is algebraic.

Proposition 6.15. *Let* $F \quad L$ *be a finite extension. Then* $F \quad L$ *is algebraic.*

Proof. An element $\alpha \; 2 \; L$ gives a chain of fields $F \quad F(\alpha) \quad L$, and then Proposition 6.13 implies that $[F(\alpha) : F] < 1$. By Proposition 6.12, α is algebraic over F. □

Definition 6.16. Let $f \; 2 \; F[x]$ be a nonconstant polynomial. Then an extension $F \quad L$ is a splitting field of f over F if

(a) $f = c(x \quad \alpha_1) \ldots (x \quad \alpha_n)$, where $c \; 2 \; F$ and $\alpha_i \; 2 \; L$, and

(b) $L = F(\alpha_1, \ldots, \alpha_n)$.

However, the previous definition tells nothing about the existence of a splitting field. So, we give the following proposition.

Proposition 6.17. *Every noncostant polynomial* $f \; 2 \; F[x]$ *has a splitting field. It is unique up to isomorphism (and thus one speaks of the splitting field of* f *over* F*).*

Proof. [5, Theorem 3.1.4, Theorem 5.1.6] □

Example 6.18. Let $f = x^4 \; 10x^2 + 1 \; 2 \; \mathbb{Q}[x]$. By the fundamental theorem of algebra $f(x)$ has 4 roots in $\mathbb{C}$ which are $\sqrt{2} + \sqrt{3}, \; \sqrt{2} \quad \sqrt{3}, \quad \sqrt{2} + \sqrt{3}, \quad \sqrt{2} \quad \sqrt{3}$. Hence the splitting field of f over $\mathbb{Q}$ is

$$\mathbb{Q}(\sqrt{2} + \sqrt{3}, \; \sqrt{2} \quad \sqrt{3}, \quad \sqrt{2} + \sqrt{3}, \quad \sqrt{2} \quad \sqrt{3}) = \mathbb{Q}(\sqrt{2}, \sqrt{3}).$$

The inclusion is obvious. The reverse inclusion follows from the fact that

$$\sqrt{2} = \frac{(\sqrt{2}+\sqrt{3})+(\sqrt{2}\quad\sqrt{3})}{2}, \qquad \sqrt{3} = \frac{(\sqrt{2}+\sqrt{3})+(\quad\sqrt{2}+\sqrt{3})}{2}.$$

We can write

$$\mathbb{Q}\quad\mathbb{Q}(\sqrt{2})\quad\mathbb{Q}(\sqrt{2},\sqrt{3}).$$

Since $x^2\quad 2$ is the minimal polynomial of $\sqrt{2}$ over $\mathbb{Q}$, then, by Proposition 6.12, $f1,\sqrt{2}g$ is a basis of $\mathbb{Q}(\sqrt{2})$ over $\mathbb{Q}$. Similarly, $f1,\sqrt{3}g$ is a basis of $\mathbb{Q}(\sqrt{2},\sqrt{3})$ over $\mathbb{Q}(\sqrt{2})$ because $x^2\sqrt{3}$ is the minimal polynomial of $\sqrt{3}$ over $\mathbb{Q}(\sqrt{2})$. So, $[\mathbb{Q}(\sqrt{2}:\mathbb{Q})]=2$ and $[\mathbb{Q}(\sqrt{2},\sqrt{3}):\mathbb{Q}(\sqrt{2})]=2$. By Proposition 6.13,

$$[\mathbb{Q}(\sqrt{2},\sqrt{3}):\mathbb{Q}] = [\mathbb{Q}(\sqrt{2},\sqrt{3}):\mathbb{Q}(\sqrt{2})][\mathbb{Q}(\sqrt{2}:\mathbb{Q})] = 2\quad 2 = 4.$$

In other words, $\mathbb{Q}\quad\mathbb{Q}(\sqrt{2},\sqrt{3})$ is a finite extension of degree 4.

Definition 6.19. An algebraic extension $F\quad L$ is normal if every irreducible polynomial in $F[x]$ that has a root in L splits completely over L.

Example 6.20. Consider $\mathbb{Q}\quad\mathbb{Q}(\sqrt[3]{2})$. The minimal polynomial of $\sqrt[3]{2}$ over $\mathbb{Q}$ is x^3+2. It has 3 roots over $\mathbb{C}$ which are $\sqrt[3]{2},\omega\sqrt[3]{2},\omega^2\sqrt[3]{2}$, where $\omega = e^{2\pi i\frac{1}{3}}$ 2 $\mathbb{C}$. Then $L = \mathbb{Q}(\sqrt[3]{2},\omega\sqrt[3]{2},\omega^2\sqrt[3]{2}) = \mathbb{Q}(\sqrt[3]{2},w)$ is the splitting field of x^3+2 over $\mathbb{Q}$. Since x^3+2 is irreducible over $\mathbb{Q}$ and the two roots $\omega\sqrt[3]{2},\omega^2\sqrt[3]{2}$ are not in $\mathbb{Q}(\sqrt[3]{2})$, the extension $\mathbb{Q}\quad\mathbb{Q}(\sqrt[3]{2})$ is not normal. However, the following proposition shows that the extension $\mathbb{Q}\quad L$ is normal.

Proposition 6.21. *A field extension $F\quad L$ is normal and finite if and only if L is the splitting field of some polynomial f 2 $F[x]$.*

Proof. [5, Theorem 5.2.4]. □

Definition 6.22. A polynomial f 2 $F[x]$ is separable if it is nonconstant and its roots in the splitting field are all simple.

In other words, f is separable if it has distinct roots.

Definition 6.23. Let $F\quad L$ be an algebraic extension.

(a) α 2 L is separable over F if its minimal polynomial over F is separable.

(b) $F\quad L$ is a separable extension if every α 2 L is separable over F.

It is well-known that f is separable if and only if $\gcd(f,f')=1$, where f' is the formal derivative of f, i.e. for $f = a_nx^n + a_{n\ 1}x^{n\ 1} +\quad + a_1x + a_0$,

$$f' = na_nx^{n\ 1} + (n\quad 1)a_{n\ 1}x^{x\ 2} +\quad + a_1.$$

From this it is easy to deduce that if F has characteristic 0, then every algebraic extension $F\quad L$ with F is separable. We give the following proposition.

Proposition 6.24. *If $F\quad L$ is an algebraic extension and F has characteristic 0, then $F\quad L$ is separable.*

Proof. Take α 2 L. Let $\mu = a_n x^n + \quad + a_1 x + a_0$ 2 $F[x]$ be its minimal polynomial over F. Let $\mu^0 = n a_n x^{n-1} + a_1$ be its formal derivative. Because F has characteristic 0, then a_n 6= 0 implies $n a_n$ 6= 0, so μ^0 has degree $n - 1$. By Proposition 6.7, μ is irreducible, so its divisors are only 1 and μ. Since $\deg(\mu^0) = n - 1$, we have $\gcd(\mu, \mu^0) = 1$, which means μ is separable. This means α is separable over F. □

The next proposition shows that every field extension $F \quad L$ which is finite and separable can be generated by one element, or that $L = F(\alpha)$ for some α 2 L. Such an element α is called a primitive element of $F \quad L$.

Proposition 6.25. *Let $F \quad L$ be a finite and separable extension. Then there is α 2 L such that $L = F(\alpha)$.*

Proof. [5, Theorem 5.4.1]. □

Example 6.26. Take the finite field extension $\mathbb{Q} \quad \mathbb{Q}(\sqrt{2}, \sqrt{3})$ from Example 6.18. By Proposition 6.24, this extension is separable, since $\mathbb{Q}$ has characteristic 0. Then, by Proposition 6.25, $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ can be generated by one element over $\mathbb{Q}$. Such an element is, for example, $\sqrt{2} + \sqrt{3}$, i.e.

$$\mathbb{Q}(\sqrt{2}, \sqrt{3}) = \mathbb{Q}(\sqrt{2} + \sqrt{3}).$$

6.2 Galois group

If L is a field, then an automorphism of L is a ring isomorphism $\sigma \colon L \to L$. We now define one of the central objects in Galois theory.

Definition 6.27. Let $F \quad L$ be a finite extension. Then $\mathrm{Gal}(L/F)$ is the set

$$\left\{\sigma \colon L \to L \mid \sigma \text{ is an automorphism, } \sigma(a) = a \text{ for all } a \in F\right\}.$$

In other words, $\mathrm{Gal}(L/F)$ consists of all automorphisms of L that are the identity on F. The basic structure of $\mathrm{Gal}(L/F)$ is that it forms a group under the operation of function composition. That's why we call $\mathrm{Gal}(L/F)$ the Galois group of $F \quad L$.

We give the following proposition.

Proposition 6.28. *Let $F \quad L$ be a finite extension and let σ 2 $\mathrm{Gal}(L/F)$. Then:*

(a) If h 2 $F[x]$ is a nonconstant polynomial with α 2 L as a root, then $\sigma(\alpha)$ is another root of h lying in L.

(b) If $L = F(\alpha_1, \ldots, \alpha_n)$, then σ is uniquely determined by its values on $\alpha_1, \ldots, \alpha_n$.

Proof. (a): Since σ preserves addition and multiplication (it is an automorphism) and, by Proposition 2.9, sends 0 to 0, we have

$$0 = \sigma(0) = \sigma(h(\alpha)) = h(\sigma(\alpha)),$$

which shows that $\sigma(\alpha)$ 2 L is another root of h.
(b): Note that every element β of $L = F(\alpha_1, \ldots, \alpha_n)$ can be written as

$$\beta = h(\alpha_1, \ldots, \alpha_n)$$

for some rational function h 2 $F(x_1, \ldots, x_n)$. Because σ preserves addition, multiplication and inverses, we have

$$\sigma(\beta) = \sigma\big(h(\alpha_1, \ldots, \alpha_n)\big) = h\big(\sigma(\alpha_1), \ldots, \sigma(\alpha_n)\big).$$

□

Example 6.29. Consider R C. Because C = R(i), then, by Proposition 6.28, any σ 2 Gal(C/R) is determined by $\sigma(i)$, since

$$\sigma(a + bi) = \sigma(a) + \sigma(b)\sigma(i) = a + b\sigma(i), \quad a, b \text{ 2 } \mathsf{R}.$$

For such an automorphism σ we can write

$$\big(\sigma(i)\big)^2 + 1 = \sigma(i^2) + \sigma(1) = \sigma(i^2 + 1) = \sigma(0) = 0 \ ) \ \sigma(i) = \quad i.$$

Let $\sigma_1(i) = i$ and $\sigma_2(i) = \quad i$. It is obvious that $\sigma_1 = \mathrm{id}_{\mathsf{C}}$ because

$$\sigma_1(a + bi) = a + b\sigma_1(i) = a + bi.$$

So, σ_1 2 Gal(C/R). To show that σ_2 is a ring homomorphism we check

$$\sigma_2\big((a + bi) + (c + di)\big) = \sigma\big((a + c) + (b + d)i\big) = (a + c) + (b + d)\sigma_2(i) =$$

$$= (a + b\sigma_2(i)) + (c + d\sigma_2(i)) = \sigma_2(a + bi) + \sigma_2(c + di),$$

$$\sigma_2\big((a + bi)(c + di)\big) = \sigma_2\big(ac \quad bd + (ad + bc)i\big) = ac \quad bd \quad (ad + bc)i =$$

$$= (a \quad bi)(c \quad di) = \sigma_2(a + bi)\sigma_2(c + di).$$

Since $\sigma_2\sigma_2 = \mathrm{id}_{\mathsf{C}}$, then σ_2 is bijective. Hence it is an automorphism of C, and so σ_2 2 Gal(C/R). Thus,

$$\mathrm{Gal}(\mathsf{C}/\mathsf{R}) = \mathsf{Z}/2\mathsf{Z}.$$

Proposition 6.30. *Let F L be a finite extension. Then its Galois group* $\mathrm{Gal}(L/F)$ *is finite.*

Proof. Since F L is finite, we can take a basis $\alpha_1, \ldots, \alpha_n$ of L over F. Then $L = F(\alpha_1, \ldots, \alpha_n)$. Let σ 2 $\mathrm{Gal}(L/F)$. By Proposition 6.28 (a), σ is uniquely determined by $\sigma(\alpha_1), \ldots, \sigma(\alpha_n)$. Let p_i be the minimal polynomial of α_i over F. Hence, by Proposition 6.28 (b), there are at most $\deg(p_i)$ possibilities for $\sigma(\alpha_i)$. That's why

$$j\mathrm{Gal}(L/F)j \quad \prod_{i=1}^{n} \deg(p_i).$$

□

Now we would like to explain how Galois groups relate to permutations of the roots of polynomials. Let $f \in F[x]$ be a separable polynomial of degree n. Let L be the splitting field of f over F. Then f splits completely over L as

$$f = c(x - \alpha_1)\dots(x - \alpha_n), \quad c \in F$$

where $\alpha_1, \dots, \alpha_n \in L$ are distinct. So, $L = F(\alpha_1, \dots, \alpha_n)$. We can define a map

$$\chi\colon \mathrm{Gal}(L/F) \to S_n \tag{6.1}$$

as follows. Given $\sigma \in \mathrm{Gal}(L/F)$, Proposition 6.28, part (a), implies that $\sigma(\alpha_i)$ is a root of f (since α_i is), so that $\sigma(\alpha_i) = \alpha_{\tau(i)}$ for some $\tau(i) \in \{1, \dots, n\}$. Note that $\tau(i)$ is uniquely determined, since $\alpha_1, \dots, \alpha_n$ are distinct. Also,

$$\tau\colon \{1, \dots, n\} \to \{1, \dots, n\}$$

is bijective since σ is. It follows that τ is a permutation, that is, $\tau \in S_n$. This defines the map (6.1).

Proposition 6.31. *The map $\chi\colon \mathrm{Gal}(L/F) \to S_n$ defined above is an injective group homomorphism.*

Proof. Suppose that $\sigma_1, \sigma_2 \in \mathrm{Gal}(L/F)$ correspond to $\tau_1, \tau_2 \in S_n$ via (6.1). This means that $\sigma_1(\alpha_i) = \alpha_{\tau_1(i)}$ for every $i = 1, \dots, n$, and similarly for σ_2 and τ_2. Then

$$(\sigma_1\sigma_2)(\alpha_i) = \sigma_1(\sigma_2(\alpha_i)) = \sigma_1(\alpha_{\tau_2(i)}) = \alpha_{\tau_1(\tau_2(i))} = \alpha_{(\tau_1\tau_2)(i)}.$$

This shows that $\sigma_1\sigma_2$ corresponds to $\tau_1\tau_2$, so that (6.1) is a group homomorphism. It remains to show that (6.1) is injective. This follows immediately from Proposition 6.28, part (b), since $L = F(\alpha_1, \dots, \alpha_n)$. □

Example 6.32. Consider the extension $\mathbb{Q} \subset L = \mathbb{Q}(\sqrt[3]{2}, w)$ from Example 6.20. Since $x^3 - 2$ is the minimal polynomial of $\sqrt[3]{2}$ over $\mathbb{Q}$ and $x^2 + x + 1$ is the minimal polynomial of w over $\mathbb{Q}$, then, by Proposition 6.28 (a),

$$\sigma \in \mathrm{Gal}(L/\mathbb{Q}) \Rightarrow \sigma(\sqrt[3]{2}) = w^i \sqrt[3]{2}, i = 0, 1, 2, \quad \sigma(w) = w^i, i = 1, 2.$$

Hence, by Proposition 6.28 (b), there are 6 different candidates for elements of $\mathrm{Gal}(L/\mathbb{Q})$. It can be verified that all of them are automorphisms of L fixing $\mathbb{Q}$ point-wise. So, $|\mathrm{Gal}(L/\mathbb{Q})| = 6$. By Proposition 6.31, $\mathrm{Gal}(L/\mathbb{Q}) = S_3$ since $|S_3| = 6$.

Definition 6.33. A field extension $F \subset L$ is a Galois extension if it is finite, normal and separable.

Example 6.34. The field extension $\mathbb{Q} \subset L$ from Example 6.20 is Galois since it is finite and normal (by Proposition 6.21) and separable (by Proposition 6.24).

Proposition 6.35. *A field extension $F \subset L$ is Galois if and only if L is a splitting field of a separable polynomial in $F[x]$.*

Proof. [5, Proposition 7.1.1]. □

It is crucial that every finite separable extension $F \subset L$ can be embedded into a larger Galois extension $F \subset N$. More precisely, we have the following result.

Proposition 6.36. *Let $F \subset L$ be a finite separable extension. Then there is an extension $L \subset N$ such that $F \subset N$ is a Galois extension.*

Proof. Since $F \subset L$ is finite and separable, by Proposition 6.25, we can write $L = F(\alpha)$ for a primitive element $\alpha \in L$. Let p be the minimal polynomial of α over F and let $\alpha_1 = \alpha, \ldots, \alpha_d$ be the roots of p. The polynomial p is separable since α is separable over F. Let N be the splitting field of p over F, i.e. $N = F(\alpha_1, \ldots, \alpha_d)$. By Proposition 6.35, $F \subset N$ is Galois. □

Definition 6.37. The field N constructed in Proposition 6.36 is called the Galois closure of L over F since N is "the smallest" extension of L which is Galois over F (see [5, Proposition 7.1.7, (b)]).

Example 6.38. Consider the finite and separable extension $\mathbb{Q} \subset \mathbb{Q}(\sqrt[3]{2})$ from Example 6.20. The minimal polynomial of $\sqrt[3]{2}$ over $\mathbb{Q}$ is $x^3 - 2$. Its splitting field over $\mathbb{Q}$ is $N = \mathbb{Q}(\sqrt[3]{2}, w)$ for $w = e^{2\pi i \frac{1}{3}}$. Thus, N is the Galois closure of $\mathbb{Q}(\sqrt[3]{2})$ over $\mathbb{Q}$.

6.3 Fundamental theorem of Galois theory

Let $F \subset L$ be a field extension and $\mathrm{Gal}(F/L)$ be its Galois group. We introduce the idea of a fixed field. Let $H \subset \mathrm{Gal}(F/L)$ be a subgroup. Then let

$$L_H = \{\alpha \in L \mid \sigma(\alpha) = \alpha \text{ for all } \sigma \in H\}.$$

It can be verified that L_H is a subfield of L which contains F. We call L_H the fixed field of H. Let K be an intermediate field between F and L, i.e.

$$F \subset K \subset L.$$

Then we can look at the elements of $\mathrm{Gal}(L/F)$ which fix K pointwise. These are the elements of $\mathrm{Gal}(L/K)$. It can be verified that $\mathrm{Gal}(L/K)$ is a subgroup of $\mathrm{Gal}(L/F)$.

Thus, to every subgroup $H \subset \mathrm{Gal}(L/F)$ we can associate an intermediate field $F \subset L_H \subset L$ and to every intermediate field $F \subset K \subset L$ we can associate a subgroup $\mathrm{Gal}(L/K) \subset \mathrm{Gal}(L/F)$. The following proposition shows that this association is a one-to-one correspondence if $F \subset L$ is a Galois extension.

Proposition 6.39. *Let $F \subset L$ be a Galois extension. Then the maps between intermediate fields $F \subset K \subset L$ and subgroups $H \subset \mathrm{Gal}(L/F)$ given by*

$$\begin{aligned} K &\mapsto \mathrm{Gal}(L/K) \\ H &\mapsto L_H \end{aligned} \tag{6.2}$$

reverse inclusions and are inverses of each other. Furthermore, $[L : K] = |\mathrm{Gal}(L/K)|$ and $[K : F] = [\mathrm{Gal}(L/F) : \mathrm{Gal}(L/K)]$. The extension $F \subset K$ is Galois if and only if $\mathrm{Gal}(L/K)$ is normal in $\mathrm{Gal}(L/F)$, and when this happens, there is an isomorphism

$$\mathrm{Gal}(K/F) = \mathrm{Gal}(L/F)/\mathrm{Gal}(L/K).$$

Proof. [5, Proposition 7.3.2]. □

Proposition 6.39 is called the fundamental theorem of Galois theory. We can deduce from it the following corollary.

Corollary 6.40. *Let $F \quad N$ be a Galois extension and let $F \quad L \quad N$ be an intermediate field. Then there is a one-to-one correspondence between intermediate fields $F \quad K \quad L$ and intermediate groups $\mathrm{Gal}(N/L) \quad H \quad \mathrm{Gal}(N/F)$. Furthermore, $[L:K] = [\mathrm{Gal}(N/K) : \mathrm{Gal}(N/L)]$ and $[K:F] = [\mathrm{Gal}(N/F) : \mathrm{Gal}(N/K)]$.*

For the needs of this work we have to generalize the second part of Proposition 6.39 in the following way.

Proposition 6.41. *Let $F \quad L$ be a Galois extension and $F \quad K \quad L$ be an intermediate field. Then there is an isomorphism*

$$\mathrm{Gal}(K/F) \overset{\kappa}{=} \mathrm{N}_{\mathrm{Gal}(L/F)}\left(\mathrm{Gal}(L/K)\right)/\mathrm{Gal}(L/K),$$

where $\mathrm{N}_G(H)$ denotes the normalizer of H in G. In particular, if $F \quad K$ is Galois, then $\mathrm{N}_{\mathrm{Gal}(L/F)}\left(\mathrm{Gal}(L/K)\right) = \mathrm{Gal}(L/F)$ and

$$\mathrm{Gal}(K/F) = \mathrm{Gal}(L/F)/\mathrm{Gal}(L/K).$$

Proof. [2, Proposition 2.6]. □

Example 6.42. Recall Example 6.32. The field extension $\mathbb{Q} \quad L = \mathbb{Q}(\sqrt[3]{2}, w)$ is Galois (as was explained in Example 6.34). Let $\sigma, \tau \in \mathrm{Gal}(L/\mathbb{Q})$ be the two generators of $\mathrm{Gal}(L/\mathbb{Q})$ defined by

$$\sigma(\sqrt[3]{2}) = w\sqrt[3]{2}, \quad \sigma(w) = w,$$
$$\tau(\sqrt[3]{2}) = \sqrt[3]{2}, \quad \tau(w) = w^2.$$

The generators σ and τ are of order 3 and 2, respectively. If we label the roots $\alpha_1 = \sqrt[3]{2}, \alpha_2 = w\sqrt[3]{2}, \alpha_3 = w^2\sqrt[3]{2}$ of $x^3 \quad 2$, then under the isomorphism $\mathrm{Gal}(L/\mathbb{Q}) = S_3$ (6.1) the automorphisms σ and τ correspond to

$$\sigma \mapsto (1\ 2\ 3), \quad \tau \mapsto (2\ 3).$$

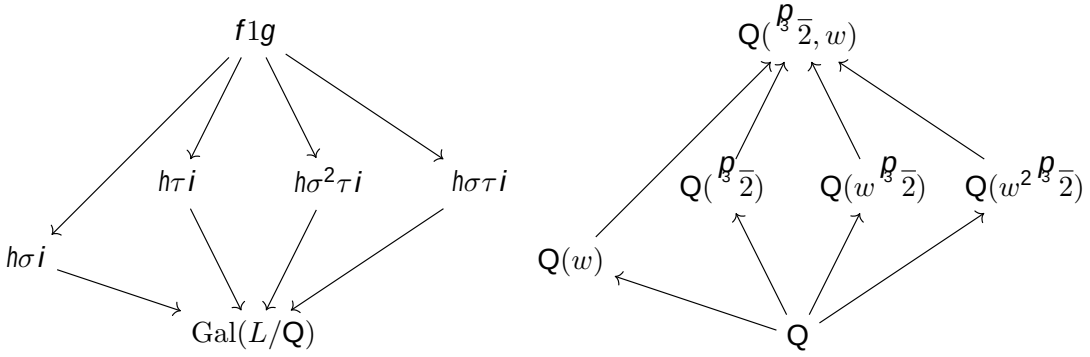


Figure 6.1: Fundamental Theorem of Galois Theory

There are 6 subgroups in S_3. Thus, by Proposition 6.39, there are 6 intermediate fields $\mathbb{Q} \subset K \subset \mathbb{Q}(\sqrt[3]{2}, w)$ (see Figure 6.1). Let's take $K = \mathbb{Q}(w)$. Then $\mathrm{Gal}(L/K) = \langle \sigma \rangle = A_3$, where A_3 denotes the alternating group on 3 elements. Since A_3 is normal in S_3, then, by Proposition 6.39, $\mathbb{Q} \subset \mathbb{Q}(w)$ is Galois and

$$\mathrm{Gal}(\mathbb{Q}(w)/\mathbb{Q}) = \mathrm{Gal}(\mathbb{Q}(\sqrt[3]{2}, w)/\mathbb{Q})/\mathrm{Gal}(\mathbb{Q}(\sqrt[3]{2}, w)/\mathbb{Q}(w)) = S_3/A_3 = S_2.$$

Another way to see that $\mathbb{Q} \subset \mathbb{Q}(w)$ is Galois (using Proposition 6.35) is to realize that $\mathbb{Q}(w)$ is the splitting field over $\mathbb{Q}$ of a separable polynomial $x^2 + x + 1$.

If we take $K = \mathbb{Q}(\sqrt[3]{2})$, then $\mathrm{Gal}(L/K) = \langle \tau \rangle = \langle (2\ 3) \rangle$. Since $\langle (2\ 3) \rangle$ is not normal in S_3, then $\mathbb{Q} \subset \mathbb{Q}(\sqrt[3]{2})$ is not Galois. However, we can apply Proposition 6.41 to $\mathbb{Q} \subset \mathbb{Q}(\sqrt[3]{2})$ and get

$$\mathrm{Gal}(\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}) = \mathrm{N}_{\mathrm{Gal}(\mathbb{Q}(\sqrt[3]{2}, w)/\mathbb{Q})}\left(\mathrm{Gal}(\mathbb{Q}(\sqrt[3]{2}, w)/\mathbb{Q}(\sqrt[3]{2}))\right) \Big/ \mathrm{Gal}(\mathbb{Q}(\sqrt[3]{2}, w)/\mathbb{Q}) =$$

$$= \mathrm{N}_{S_3}\left(\langle (2\ 3) \rangle\right) / \langle (2\ 3) \rangle = \langle (2\ 3) \rangle / \langle (2\ 3) \rangle = \{1\}.$$

Finally, we would like to give a useful remark which we are going to use in Chapter 8.

Remark 6.43. Let F be a field of characteristic 0 and let $F \subset L$ be a finite extension of degree d. By Proposition 6.15, $F \subset L$ is algebraic and, by Proposition 6.24, $F \subset L$ is separable. By Proposition 6.25, there is an element $\alpha \in L$ such that $L = F(\alpha)$. By Proposition 6.12, its minimal polynomial μ over F has degree d. Denote the roots of μ as $\alpha = \alpha_1, \ldots, \alpha_d$. They are all distinct since μ is separable. Hence, according to Proposition 6.36, the Galois closure of L over F is the splitting field of μ over F which is

$$N = F(\alpha_1, \ldots, \alpha_d).$$

Let $G = \mathrm{Gal}(N/F)$. By the fundamental theorem of Galois theory (Proposition 6.39), there is a one-to-one correspondence between the intermediate fields $F \subset K \subset N$ and the subgroups $H \subset G$. Moreover, by Corollary 6.40, there is a one-to-one correspondence between intermediate fields

$$F \subset K \subset L$$

and intermediate groups

$$\mathrm{Gal}(N/L) \subset H \subset G.$$

In particular, if there is no intermediate group between $\mathrm{Gal}(N/L)$ and G, then there is no intermediate field between F and L.

7 Branched Covers of Algebraic Varieties

Solving systems of polynomial equations is a common thing in practice (for example, in computer vision). Every formulation of the problem by polynomial equations $G(\mathbf{x}, \mathbf{p}) = \mathbf{0}$ contains the unknowns $\mathbf{x}$ 2 C^n and parameters $\mathbf{p}$ 2 C^m. We are usually interested in finding the solutions (i.e. determining the unknowns) given the certain values of parameters. If for a generic choice of parameters $\mathbf{p}_0$ 2 C^m the system has $d < 1$ solutions, then we can create a map

$$\begin{aligned} f \colon \mathbf{V}(G(\mathbf{x}, \mathbf{p})) \ &! \ \mathsf{C}^m \\ \begin{bmatrix} \mathbf{x} \\ \mathbf{p} \end{bmatrix} &7! \ \mathbf{p} \end{aligned} \tag{7.1}$$

which projects the space of the unknowns and parameters, satisfying G, to the space of parameters. This projection is a finite map of degree d between a□ne varieties.

The main purpose of this chapter is to establish the connections between algebraic geometry, algebraic topology and Galois theory. We state the result that a finite rational map of degree d between a□ne varieties can be understood as a branched cover (meaning it is a covering map except on a small set) of degree d in the analytic topology. When it comes to revealing the properties of this branched cover we have to use results coming from another branch of mathematics, called analytic geometry. There is a strong connection between algebraic geometry and analytic geometry which is described by the paper "Géométrie algébrique et géométrie analytique" written by J.P. Serre. The results from this paper are usually referred to as "GAGA". In this chapter we give one of the important GAGA results which states that an irreducible a□ne variety without its singular points is connected in the analytic topology. We don't explain any proofs here since most of them we don't understand completely.

We describe here the main object of this work, namely the Galois/monodromy group of a branched cover. This is where we will combine our knowledge from Chapters 4, 5 and 6. Finally, we describe the notion of symmetries of branched covers and show how they can be found.

7.1 Finite rational maps

We will start by stating some of GAGA results. Let X C^n be an a□ne variety given by $\mathbf{F} = \left[f_1, \ldots, f_r\right]^>$, f_i 2 C$[x_1, \ldots, x_n]$. Let $\mathrm{Sing}(X)$ X be the set of singular points of X, i.e. the points p 2 X for which the r n Jacobian matrix

$$J(\mathbf{F}) = \left[\frac{\partial \mathbf{F}}{\partial x_1} \quad \ldots \quad \frac{\partial \mathbf{F}}{\partial x_n}\right],$$

evaluated at p, has rank less than $n \quad \dim(X)$, where $\dim(X)$ denotes the dimension of X [6, Chapter 9]. The set $\mathrm{Sing}(X)$ is a subvariety of X since it is given by $\mathbf{F}$ together with $(n \quad \dim(X)) \quad (n \quad \dim(X))$ minors of $J(\mathbf{F})$. We denote $X_{\text{an}} = X n \mathrm{Sing}(X)$ the set of non-singular points in X and we give this set the subspace topology induced from the standard topology on $\mathbb{C}^n$. This topology on X_{an} is called the analytic topology. We give the first GAGA result.

Proposition 7.1. *Let $X \quad \mathbb{C}^n$ be an a ne variety. Then X_{an} is a complex manifold.*

Proof. [22, Lemma 2.2]. □

Corollary 7.2. *Let $X \quad \mathbb{C}^n$ be an a ne variety. Then X_{an} is locally path-connected.*

Proof. This is because every complex manifold, by definition, looks locally like an open unit disk in $\mathbb{C}^m$, which is path-connected. □

Example 7.3. Let $X = \mathbf{V}(x^2 \quad y^2) \quad \mathbb{C}^2$. Then $\dim(X) = 1$. Denote $\mathbf{F} = \begin{bmatrix} x^2 & y^2 \end{bmatrix}$. Then

$$J(\mathbf{F}) = \begin{bmatrix} 2x & 2y \end{bmatrix}.$$

Hence

$$\mathrm{Sing}(X) = \mathbf{V}(x^2 \quad y^2, 2x, 2y) = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}.$$

Then

$$X_{\text{an}} = X n \mathrm{Sing}(X) = \mathbf{V}(x^2 \quad y^2) n \, f0g \, .$$

By Proposition 7.1, X_{an} is a complex manifold. By Corollary 7.2, it is locally path-connected.

We state the second GAGA result.

Proposition 7.4. *Let $X \quad \mathbb{C}^n$ be an irreducible a ne variety. Then X_{an} is connected.*

Proof. [28, Theorem 8.3]. □

Corollary 7.5. *Let $X \quad \mathbb{C}^n$ be an irreducible a ne variety. Then X_{an} is path-connected.*

Proof. Follows from Proposition 2.41. □

Example 7.6. Let $X = \mathbf{V}(y^2 \quad x^2 \quad x^3) \quad \mathbb{C}^2$. This variety is irreducible, since $y^2 \quad x^2 \quad x^3$ is irreducible over $\mathbb{C}$. The singular points of X are

$$\mathrm{Sing}(X) = \mathbf{V}(y^2 \quad x^2 \quad x^3, 2y, \quad 2x \quad 3x^2) = f0g.$$

Then $X_{\text{an}} = \mathbf{V}(y^2 \quad x^2 \quad x^3) n f0g$ is a path-connected complex manifold (however, you cannot see it from the picture of X_{an} over the real numbers).

The next proposition shows that every dominant rational map between irreducible a□ne varieties of the same dimension is a covering map in the analytic topology.

Proposition 7.7. *Let $f\colon X \dashrightarrow Z$ be a dominant rational map between irreducible affine varieties of the same dimension. Then there is a Zariski dense subset $U \subseteq Z$ containing a Zariski open subset $V \subseteq Z$ such that the map*

$$f\big|_{f^{-1}(U)}\colon f^{-1}(U) \to U.$$

is a covering map in the analytic topology. Moreover, by Corollaries 7.2 and 7.5, $f^{-1}(U)$ (and hence U) is path-connected and locally path-connected in the analytic topology. We call f the branched cover, meaning it is almost a covering map except on a small set $Z \setminus U$. We denote $f\big|_{f^{-1}(U)}$ by f_c and $f^{-1}(U)$ by $\widetilde{U}$. Since $\widetilde{U}$ is path-connected, then U is connected. If we let $d = |f_c^{-1}(u_0)|$ for some $u_0 \in U$, then by Remark 5.11, $|f_c^{-1}(u)| = d$ for all $u \in U$. We call d the degree of the branched cover f.

Proof. [22, Chapter 2]. □

We explain how the set U can be constructed when $Z = \mathbb{C}^m$ and $f\colon X \dashrightarrow Z$ is the projection:

$$\begin{aligned} f\colon X &\to Z \\ \begin{bmatrix} \mathbf{x} \\ \mathbf{z} \end{bmatrix} &\mapsto \mathbf{z} \end{aligned}$$

where $\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^\top$ and $\mathbf{z} = \begin{bmatrix} z_1 & \dots & z_m \end{bmatrix}^\top$. Let $\mathbf{F} = \begin{bmatrix} f_1 & \dots & f_r \end{bmatrix}^\top$ be the column vector of polynomials which define X. Then the Jacobian of $\mathbf{F}$ is the following polynomial matrix

$$J(\mathbf{F}) = \begin{bmatrix} \frac{\partial \mathbf{F}}{\partial x_1} & \dots & \frac{\partial \mathbf{F}}{\partial x_n} & \frac{\partial \mathbf{F}}{\partial z_1} & \dots & \frac{\partial \mathbf{F}}{\partial z_m} \end{bmatrix}$$

of size $r \times (n+m)$. Let $C_f \subseteq X$ be the set of critical points of f, i.e. the singular points of X or the nonsingular points $x \in X$ where the differential of f

$$\mathrm{d}_x f\colon T_x X \to T_{f(x)} Z$$

fails to surject. This is exactly the set of points $p \in X$ for which the matrix

$$J_{\mathbf{x}}(\mathbf{F})_p \overset{\text{def}}{=} \left[\frac{\partial \mathbf{F}}{\partial x_1} & \dots & \frac{\partial \mathbf{F}}{\partial x_n} \right]_p = \left[\frac{\partial \mathbf{F}}{\partial x_1}(p) & \dots & \frac{\partial \mathbf{F}}{\partial x_n}(p) \right] \in \mathbb{C}^{r \times n}$$

has rank less than n [16, Proposition 3.4]. Then $U = f(X) \setminus f(C_f)$.

In Figure 7.1 you can see a formal example of a branched cover. The points p_1 and p_2 are the critical points of f: p_1 is singular (there is no tangent space to X at p_1) and the differential

$$\mathrm{d}_{p_2} f\colon T_{p_2} X \to T_{f(p_2)} Z$$

maps $T_{p_2} X$ to a point $f(p_2)$ (since $T_{p_2} X$ is "perpendicular" to $T_{f(p_2)} Z$), which implies $\mathrm{d}_{p_2} f$ is not surjective. For the points u from $U = f(X) \setminus \{f(p_1), f(p_2)\}$ the fiber $f^{-1}(u)$ consists of 2 points. By Remark 5.11, the cardinality $|f^{-1}(u)|$ has to be the same for all $u \in U$.

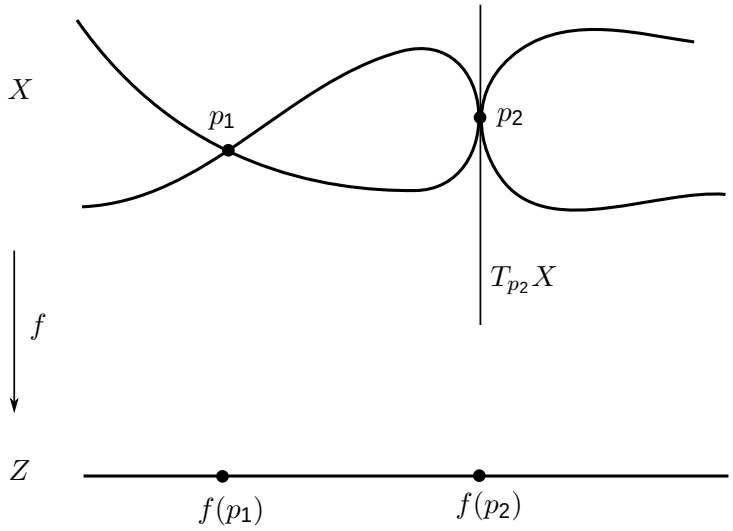


Figure 7.1

Example 7.8. Let $X = \mathbf{V}(x^3 \quad z) \quad \mathbb{C}^2$ and $Z = \mathbb{C}$. Then let

$$f\colon X \dashrightarrow Z$$
$$\begin{bmatrix} x \\ z \end{bmatrix} \mapsto z$$

This map is defined everywhere on X. It is surjective by the Fundamental Theorem of Algebra. Also X and Z are irreducible since $\mathbf{I}(X) = \langle x^3 \quad z \rangle$ and $\mathbf{I}(Z) = \langle 0 \rangle$ are prime ideals. Also $\dim(X) = \dim(Z) = 1$. Then we can apply Proposition 7.7. The Jacobian of $\mathbf{F} = \begin{bmatrix} x^3 & z \end{bmatrix}$ equals

$$J(\mathbf{F}) = \begin{bmatrix} \frac{\partial \mathbf{F}}{\partial x} & \frac{\partial \mathbf{F}}{\partial z} \end{bmatrix} = \begin{bmatrix} 3x^2 & 1 \end{bmatrix}.$$

Let $p = \begin{bmatrix} x_0 & z_0 \end{bmatrix}^{>} \in X$. Then

$$J_x(\mathbf{F})_p = \begin{bmatrix} 3x_0^2 \end{bmatrix}.$$

This matrix has rank less than 1 if and only if $x_0 = 0$. The only point in X with $x_0 = 0$ is the point $p = \begin{bmatrix} 0 & 0 \end{bmatrix}^{>} \in X$. This is the point where the differential $\mathrm{d}_p f$ fails to surject. Thus, $C_f = \{\begin{bmatrix} 0 & 0 \end{bmatrix}^{>}\}$ is the set of critical points of f. Since f is surjective, we have $U = f(X) \setminus f(C_f) = \mathbb{C} \setminus \{0\}$ and

$$f_c\colon \widetilde{U} \to U$$

is a covering map in the analytic topology. Let $u_0 = 1 \in U$. Then

$$f_c^{-1}(u_0) = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} e^{2\pi i \frac{1}{3}} \\ 1 \end{bmatrix}, \begin{bmatrix} e^{2\pi i \frac{2}{3}} \\ 1 \end{bmatrix} \right\}.$$

Since $|f_c^{-1}(u_0)| = 3$, then, by Remark 5.11, $|f_c^{-1}(u)| = 3$ for all $u \in U$.

Definition 7.9. A dominant rational map $f\colon X \dashrightarrow Z$ between irreducible affine varieties is said to be finite of degree d if there is a Zariski dense open subset $V \subseteq Z$ such that for every $z \in V$ the fiber $f^{-1}(z)$ consists of d distinct points.

Proposition 7.10. *A dominant rational map $f\colon X \dashrightarrow Z$ is finite of degree d if and only if it induces a finite field extension $f\ (\mathsf{C}(Z)) \quad \mathsf{C}(X)$ of degree d.*

Proof. [14, Proposition 7.16]. □

Example 7.11. Take f from Example 7.8. We saw that for $U = \mathsf{C}nf0g$ the map

$$f_c\colon \widetilde{U} \ ! \ U$$

is a covering map of degree 3 in the analytic topology. By Definition 7.9, f is a finite rational map of degree 3. Thus, by Proposition 7.10, f induces an inclusion of function fields

$$f \ : \mathsf{C}(Z) \ ! \ \mathsf{C}(X)$$

of degree 3.

7.2 Galois/monodromy group

In this section we explain how the monodromy group of f_c can be obtained as the Galois group of a field extension. Let $f\colon X \dashrightarrow Z$ be a branched cover of degree d. Then, by Proposition 7.10, it induces the inclusion $f \ : \mathsf{C}(Z) \hookrightarrow \mathsf{C}(X)$ of function fields of degree d. Then the certain restriction $f_c\colon \widetilde{U} \ ! \ U$ of f is a covering map of degree d in the analytic topology. Since C is the field of characteristic zero, then so are $\mathsf{C}(Z)$ and $\mathsf{C}(X)$. Hence we can apply what was said in Remark 6.43 to the field extension $\mathsf{C}(Z) \quad \mathsf{C}(X)$. Let $\alpha\ 2\ \mathsf{C}(X)$ be a primitive element of this field extension. Then $\mathsf{C}(X) = \mathsf{C}(Z)(\alpha)$. Let μ be the minimal polynomial of α over $\mathsf{C}(Z)$ and let $\alpha_1 = \alpha, \ldots, \alpha_d$ be its roots in the splitting field. Denote $R = f\alpha_1, \ldots, \alpha_d g$ and let $N_f = \mathsf{C}(Z)(\alpha_1, \ldots, \alpha_d)$ be the Galois closure of $\mathsf{C}(Z) \quad \mathsf{C}(X)$. Denote $\mathtt{Gal}(f) = \mathrm{Gal}(N_f/\mathsf{C}(Z))$. We will say that $\mathtt{Gal}(f)$ is the Galois group of the branched cover f. Then there is an injective group isomorphism $\chi\colon \mathtt{Gal}(f) \hookrightarrow \mathsf{S}(R)$ given by 6.1. Let $u\ 2\ U$ and $F = f_c\ ^1(u)$. We give the following proposition.

Proposition 7.12. *The Galois group $\mathtt{Gal}(f)$ and the monodromy group $\mathtt{Mon}_F(f_c)$ are isomorphic as permutation groups.*

Proof. [13, Section I]. □

In other words, there is a bijection $\eta\colon R \ ! \ F$ and a group isomorphism $\lambda\colon \mathtt{Gal}(f) \ ! \ \mathtt{Mon}_F(f_c)$ such that for $\phi(\sigma) = \eta\sigma\eta\ ^1$ and the identity map id_M on $\mathtt{Mon}_F(f_c)$ the diagram

$$\begin{array}{ccc} \mathtt{Gal}(f) & \xhookrightarrow{\ \chi\ } & \mathsf{S}(R) \\ {\scriptstyle =}\downarrow{\scriptstyle \lambda} & & {\scriptstyle =}\downarrow{\scriptstyle \phi} \\ \mathtt{Mon}_F(f_c) & \xhookrightarrow{\ \mathrm{id}_M\ } & \mathsf{S}(F) \end{array} \tag{7.2}$$

commutes, i.e. $\mathrm{id}_M\lambda = \phi\chi$.

We give the following example which was created with the help of Tim Du□.

Example 7.13. Let $f\colon X \to Z$ be the map from Example 7.8, i.e. $X = \mathbf{V}(x^3 - z) \subset \mathbb{C}^2$, $Z = \mathbb{C}$ and

$$\begin{aligned} f\colon X &\to Z \\ \begin{bmatrix} x \\ z \end{bmatrix} &\mapsto z \end{aligned}$$

It was shown that for $U = \mathbb{C}\setminus\{0\}$, the restriction of f to $\widetilde{U} = f^{-1}(U)$

$$f_c\colon \widetilde{U} \to U$$

is a covering map of degree 3 in the analytic topology. Also the field extension $\mathbb{C}(Z) \subset \mathbb{C}(X)$ has degree 3. Let $u = 1 \in U$ and

$$\widetilde{u}_1 = \begin{bmatrix} e^{2\pi i\frac{1}{3}} \\ 1 \end{bmatrix}, \widetilde{u}_2 = \begin{bmatrix} e^{2\pi i\frac{2}{3}} \\ 1 \end{bmatrix}, \widetilde{u}_3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad F = f^{-1}(u) = \{\widetilde{u}_1, \widetilde{u}_2, \widetilde{u}_3\}.$$

We now determine the monodromy group $\mathrm{Mon}_F(f_c)$. Let's take a look at the fundamental group $\pi_1(U, u)$. Since U looks like a plane with one point $0 \in \mathbb{C}$ removed, then a loop γ around 0 based at u represents the generator of $\pi_1(U, u)$ (as it was explained in the end of Section 5.1). In other words, $\pi_1(U, u) = \langle[\gamma]\rangle = \mathbb{Z}$. Let's take a loop

$$\begin{aligned} \gamma\colon I &\to U \\ t &\mapsto e^{2\pi it} \end{aligned}$$

based at u (since $\gamma(0) = 1 = \gamma(1)$). It obviously encircles $0 \in \mathbb{C}$. By the path lifting property (see the paragraph after Proposition 5.13), this loop has 3 unique lifts $\widetilde{\gamma}_1, \widetilde{\gamma}_2, \widetilde{\gamma}_3$ starting at

$$\widetilde{u}_1 = \begin{bmatrix} e^{2\pi i\frac{1}{3}} \\ 1 \end{bmatrix}, \widetilde{u}_2 = \begin{bmatrix} e^{2\pi i\frac{2}{3}} \\ 1 \end{bmatrix}, \widetilde{u}_3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

respectively. These lifts have the following form:

$$\begin{aligned} \widetilde{\gamma}_j\colon I &\to \widetilde{U} \\ t &\mapsto \begin{bmatrix} e^{2\pi i\frac{j+t}{3}} \\ e^{2\pi it} \end{bmatrix} \end{aligned}$$

since $f_c\widetilde{\gamma}_j = \gamma$ for all $j = 1, 2, 3$. The monodromy group $\mathrm{Mon}_F(f_c)$ is the image of the group homomorphism (5.1). Since $[\gamma\ \ ] = [\gamma]^{-1}$ generates $\pi_1(U, u)$, then σ_γ generates $\mathrm{Mon}_F(f_c)$. It is obvious that

$$\begin{aligned} \sigma_\gamma(\widetilde{u}_1) &= \widetilde{\gamma}_1(1) = \widetilde{u}_2, \\ \sigma_\gamma(\widetilde{u}_2) &= \widetilde{\gamma}_2(1) = \widetilde{u}_3, \\ \sigma_\gamma(\widetilde{u}_3) &= \widetilde{\gamma}_3(1) = \widetilde{u}_1. \end{aligned}$$

Hence

$$\mathrm{Mon}_F(f_c) = \langle\sigma_\gamma\rangle = \langle(1\ 2\ 3)\rangle = \mathbb{Z}/3\mathbb{Z}.$$

Let N_f be the Galois closure of $\mathbb{C}(Z) \quad \mathbb{C}(X)$, i.e. we have a chain of fields

$$\mathbb{C}(Z) \quad \mathbb{C}(X) \quad N_f.$$

Then, by Proposition 7.12,

$$\mathrm{Gal}(N_f/\mathbb{C}(Z)) = \mathrm{Mon}_F(f_c) = \mathbb{Z}/3\mathbb{Z}.$$

From Galois theory it follows that $F \quad L$ is a Galois extension if and only if $j\mathrm{Gal}(L/F)j = [L : F]$ (see [5, Proposition 7.1.5]). Since $\mathbb{C}(Z) \quad N_f$ is a Galois extension, then

$$[N_f : \mathbb{C}(Z)] = j\mathrm{Gal}(N_f/\mathbb{C}(Z))j = 3.$$

However, we know that $[\mathbb{C}(X) : \mathbb{C}(Z)] = 3$. By Proposition 6.13,

$$[N_f : \mathbb{C}(Z)] = [N_f : \mathbb{C}(X)][\mathbb{C}(X) : \mathbb{C}(Z)],$$

which implies

$$[N_f : \mathbb{C}(X)] = 1 \quad ( ) \quad N_f = \mathbb{C}(X).$$

In other words, the field extension $\mathbb{C}(Z) \quad \mathbb{C}(X)$ is Galois.

More generally we obtain the following proposition.

Proposition 7.14. *If $f \colon X \dashrightarrow Z$ is a branched cover of degree d such that $j\mathrm{Mon}_F(f_c)j = d$, then the field extension $\mathbb{C}(Z) \quad \mathbb{C}(X)$ is Galois and*

$$\mathrm{Gal}(\mathbb{C}(X)/\mathbb{C}(Z)) \overset{\lambda}{=} \mathrm{Mon}_F(f_c)$$

for λ from (7.2).

Proof. Let N_f be the Galois closure of $\mathbb{C}(Z) \quad \mathbb{C}(X)$. Then, by Proposition 7.12,

$$j\mathrm{Gal}(N_f/\mathbb{C}(Z))j = j\mathrm{Mon}_F(f_c)j = d.$$

From Galois theory it follows that $F \quad L$ is a Galois extension if and only if $j\mathrm{Gal}(L/F)j = [L : F]$ (see [5, Proposition 7.1.5]). Since $\mathbb{C}(Z) \quad N_f$ is a Galois extension, then

$$[N_f : \mathbb{C}(Z)] = j\mathrm{Gal}(N_f/\mathbb{C}(Z))j = d.$$

However, by Proposition 7.10, we know that $[\mathbb{C}(X) : \mathbb{C}(Z)] = d$ since f has degree d. By Proposition 6.13,

$$[N_f : \mathbb{C}(Z)] = [N_f : \mathbb{C}(X)][\mathbb{C}(X) : \mathbb{C}(Z)],$$

which implies

$$[N_f : \mathbb{C}(X)] = 1 \quad ( ) \quad N_f = \mathbb{C}(X).$$

□

Definition 7.15. In the situation of Proposition 7.14 we say that the branched cover f is Galois.

We show another example of a branched cover which is not Galois.

Example 7.16. Let $X = \mathbf{V}(x^6 + 2x^4 + 3x^2 + z) \quad \mathbb{C}^2$ and $Z = \mathbb{C}$. We define a map

$$f\colon X \dashrightarrow Z$$
$$\begin{bmatrix} x \\ z \end{bmatrix} \mapsto z$$

This map is defined everywhere on X. It is surjective by the Fundamental Theorem of Algebra. Also X and Z are irreducible since $\mathbf{I}(X) = \langle x^6 + 2x^4 + 3x^2 + z \rangle$ and $\mathbf{I}(Z) = \langle 0 \rangle$ are prime ideals. Also $\dim(X) = \dim(Z) = 1$. The Jacobian of $\mathbf{F}$ equals

$$J(\mathbf{F}) = \begin{bmatrix} 6x^5 + 8x^3 + 6x^2 & 1 \end{bmatrix}.$$

Let $p = \begin{bmatrix} x_0 & z_0 \end{bmatrix}^{>} \in X$. Then

$$J_x(\mathbf{F})_p = \begin{bmatrix} 6x_0^5 + 8x_0^3 + 6x_0^2 \end{bmatrix}.$$

This matrix has rank less than 1 if and only if x_0 is the root of $6x^5 + 8x^3 + 6x^2$. The image of the critical points of f are the roots of the generator of the elimination ideal

$$\langle x^6 + 2x^4 + 3x^2 + z, 6x^5 + 8x^3 + 6x^2 \rangle \cap \mathbb{C}[z] = \langle 27z^3 \quad 76z^2 + 72z \rangle.$$

Thus, $U = f(X) \setminus f(C_f) = Z \setminus \mathbf{V}(27z^3 \quad 76z^2 + 72z)$ and

$$f_c\colon \widetilde{U} \to U$$

is a covering map in the analytic topology of degree 6. Let $u \in U$. The fundamental group of U is isomorphic to the free product

$$\pi_1(U, u) = \mathbb{Z} \quad \mathbb{Z} \quad \mathbb{Z},$$

since we removed 3 points from $\mathbb{C}$ to get U. Let $F = f_c^{-1}(u)$. Then the monodromy group $\mathrm{Mon}_F(f_c)$ is the image of $\pi_1(U, u)$ under the group homomorphism (5.1). In this example we don't construct loops and their lifts explicitly as it was done in Example 7.13. The monodromy group $\mathrm{Mon}_F(f_c)$ can be computed numerically using [16]. In this case we get

$$\mathrm{Mon}_F(f_c) = \Big\langle (1\ 4), (5\ 1)(6\ 4), (2\ 4)(1\ 3) \Big\rangle.$$

The 3 permutations on the right are the monodromy permutations σ_γ for loops γ based at u and encircling each of the 3 points from $f(C_f) \quad Z$. These 3 loops generate $\pi_1(U, u)$. Since $|\mathrm{Mon}_F(f_c)| = 48 \neq 6 = [\mathbb{C}(X) : \mathbb{C}(Z)]$, then f is not Galois.

7.3 Symmetries of branched covers

Let $f\colon X \dashrightarrow Z$ be a branched cover. In this section we will explain what do we mean by symmetries of branched covers. Recall the notion of a birational map from Definition 4.45.

Definition 7.17. A birational automorphism of X is a birational map $\varphi\colon X \dashrightarrow X$.

Remark 7.18. We are not interested in all birational automorphisms of X, but only in some special. From practical point of view, if we have a branched cover $f\colon X \dashrightarrow Z$ as in (7.1), then for every instance $z \in Z$ of the problem its solutions are the points in the fiber $f^{-1}(z)$. We would like to exploit the global symmetry of the whole problem in such a way that we are able to restrict this symmetry to the fiber $f^{-1}(z)$ of every instance $z \in Z$.

Thus, the only reasonable way to define the concept of symmetries is the following one.

Definition 7.19. Let $f\colon X \dashrightarrow Z$ be a branched cover. We define the set of symmetries of f to be the set of all birational automorphisms φ of X such that $f = f\varphi$ (where defined). We denote this set by $\mathtt{Bir}(f)$.

Such a definition puts a certain restriction on $\varphi \in \mathtt{Bir}(f)$: it preserves the fiber $f^{-1}(z)$ for every $z \in Z$. (Notice the analogy with Definition 5.18.) It is easy to prove that for every branched cover $f\colon X \dashrightarrow Z$ its set of symmetries forms a group.

Proposition 7.20. *Let $f\colon X \dashrightarrow Z$ be a branched cover. Then $\mathtt{Bir}(f)$ is a group under the operation of function composition.*

Proof. The operation of function composition is associative. If $\varphi_1, \varphi_2 \in \mathtt{Bir}(f)$, then

$$f(\varphi_1\varphi_2) = (f\varphi_1)\varphi_2 = f\varphi_2 = f \Longrightarrow \varphi_1\varphi_2 \in \mathtt{Bir}(f).$$

The identity map $\mathrm{id}_X \in \mathtt{Bir}(f)$ since, obviously, $f = f\mathrm{id}_X$. If $f = f\varphi$, then

$$f\varphi^{-1} = (f\varphi)\varphi^{-1} = f(\varphi\varphi^{-1}) = f\mathrm{id}_X = f \Longrightarrow \varphi^{-1} \in \mathtt{Bir}(f).$$

□

It is crucial that the group $\mathtt{Bir}(f)$ can be obtained as the Galois group of a certain field extension.

Proposition 7.21. *Let $f\colon X \dashrightarrow Z$ be a branched cover with its induced embedding of function fields $f^* : \mathbb{C}(Z) \hookrightarrow \mathbb{C}(X)$. Then*

$$\mathtt{Bir}(f) \overset{\beta}{=} \mathrm{Gal}\Big(\mathbb{C}(X)/f^*(\mathbb{C}(Z))\Big).$$

Proof. We define β to be

$$\begin{aligned} \beta\colon \mathtt{Bir}(f) &\to \mathrm{Gal}\Big(\mathbb{C}(X)/f^*(\mathbb{C}(Z))\Big) \\ \varphi &\mapsto (\varphi^{-1})^* \end{aligned}$$

We at first verify that β is well-defined. By Proposition 4.46, $(\varphi^{-1})^*$ is a ring isomorphism from $\mathbb{C}(X)$ to $\mathbb{C}(X)$, so it is an automorphism of $\mathbb{C}(X)$. According to Definition 7.19, we have $f = f\varphi^{-1}$. By Proposition 4.44, $f^* = (\varphi^{-1})^* f^*$. But this equality of maps exactly means that $f^*(\chi) = (\varphi^{-1})^*(f^*(\chi))$ for every $\chi \in \mathbb{C}(Z)$, or that $(\varphi^{-1})^*$ fixes $f^*(\mathbb{C}(Z))$ pointwise. This exactly means that $(\varphi^{-1})^* \in \mathrm{Gal}\Big(\mathbb{C}(X)/f^*(\mathbb{C}(Z))\Big)$.

Now we show that β is a group homomorphism. Take $\varphi_1, \varphi_2 \in \mathtt{Bir}(f)$. Then using Proposition 4.44 we get

$$\beta(\varphi_1\varphi_2) = ((\varphi_1\varphi_2)^{-1}) = (\varphi_2^{-1}\varphi_1^{-1}) = (\varphi_1^{-1})\,(\varphi_2^{-1}) = \beta(\varphi_1)\beta(\varphi_2).$$

We claim that β is bijective. By Proposition 4.46, the function $\varphi \mapsto (\varphi^{-1})$ is injective on the set of all birational maps from X to X. So, it is injective when restricted to $\mathtt{Bir}(f)$. To prove that β is surjective, take any element from $\mathrm{Gal}\Big(\mathsf{C}(X)/f^{*}(\mathsf{C}(Z))\Big)$. It is an automorphism of $\mathsf{C}(X)$, and so, by Proposition 4.46, it has the form (φ^{-1}) for some birational map $\varphi\colon X \dashrightarrow X$. We also know that $f = (\varphi^{-1})^{*} f$ because (φ^{-1}) fixes $f^{*}(\mathsf{C}(Z))$ pointwise. By Proposition 4.44, we have $f = (\varphi^{-1})^{*} f = (f\varphi^{-1})$. Because the operator $\varphi \mapsto (\varphi^{-1})$ is injective, then $f = f\varphi^{-1}$. Thus, $\varphi \in \mathtt{Bir}(f)$.

We proved that β is a bijective group homomorphism. Thus, β is an isomorphism.

□

An immediate consequence of Proposition 7.21 is that $\mathtt{Bir}(f)$ is a finite group, since the field extension $\mathsf{C}(Z) \subset \mathsf{C}(X)$ is finite (see Proposition 6.30). Our aim now is to reveal the connection between $\mathtt{Bir}(f)$ and $\mathtt{Gal}(f)$. We have the chain of fields

$$\mathsf{C}(Z) \subset \mathsf{C}(X) \subset N_f.$$

where the field extension $\mathsf{C}(Z) \subset N_f$ is Galois (first paragraph in Section 7.2). Let $G = \mathrm{Gal}(N_f/\mathsf{C}(X))$ be the subgroup of $\mathtt{Gal}(f)$. Hence by Proposition 6.41, there is an isomorphism

$$\mathtt{Bir}(f) \overset{\kappa\beta}{=} \mathrm{N}_{\mathtt{Gal}(f)}(G)/G.$$

Let's look at the image of G by χ. This is a permutation group $\chi(G) \subset \mathsf{S}(R)$ isomorphic to G. Recall that $\mathsf{C}(X) = \mathsf{C}(Z)(\alpha_1)$. Then G consists of elements of $\mathtt{Gal}(f)$ which fix α_1. Hence $\chi(G) = \mathrm{Stab}_{\chi(\mathtt{Gal}(f))}(\alpha_1)$. It is also easy to see that

$$\mathrm{N}_{\mathtt{Gal}(f)}(G)/G \overset{\widetilde{\chi}}{=} \mathrm{N}_{\chi(\mathtt{Gal}(f))}\big(\chi(G)\big)/\chi(G).$$

since χ is a group homomorphism. Then relation (3.4) shows that

$$\mathrm{N}_{\chi(\mathtt{Gal}(f))}\big(\chi(G)\big)/\chi(G) \overset{\varrho}{=} \mathrm{C}_{\mathsf{S}(R)}\Big(\chi(\mathtt{Gal}(f))\Big).$$

Combining $\beta, \kappa, \widetilde{\chi}$ and ϱ together we obtain

$$\mathtt{Bir}(f) \overset{\varrho\widetilde{\chi}\kappa\beta}{=} \mathrm{C}_{\mathsf{S}(R)}\Big(\chi(\mathtt{Gal}(f))\Big) \tag{7.3}$$

We are now able to draw the following diagram:

$$\begin{array}{ccccccc}
\mathtt{Gal}(f) & \xhookrightarrow{\chi} & \mathsf{S}(R) & \xhookleftarrow{\qquad\varrho\widetilde{\chi}\kappa\beta\qquad} & \mathtt{Bir}(f) \\
{\scriptstyle =}\big\downarrow{\scriptstyle \lambda} & & {\scriptstyle =}\big\downarrow{\scriptstyle \phi} & & \\
\mathtt{Mon}_F(f_c) & \xhookrightarrow{\mathrm{id}_M} & \mathsf{S}(F) & \xhookleftarrow{\mathrm{id}_D} & \mathtt{Deck}_F(f_c) & \xleftarrow[=]{\psi} & \mathtt{Deck}(f_c)
\end{array} \tag{7.4}$$

where ψ is an isomorphism (5.4) and id_D is the identity map on $\mathtt{Deck}_F(f_c)$. Diagram (7.4) is commutative since diagram (7.2) is commutative. Since, by Proposition 7.7, $f_c\colon \widetilde{U} \to U$ is path-connected and locally path-connected covering space, then, by Propositions 5.26 and 5.27, we conclude that

$$\mathtt{Bir}(f) = \mathtt{Deck}(f_c) \tag{7.5}$$

via the isomorphism $\psi = {}^{1}\phi\varrho\widetilde{\chi}\kappa\beta$.

Let $\varphi \in \mathtt{Bir}(f)$. Since φ is birational and $f\colon X \dashrightarrow Z$ is finite of degree d, then a rational map $f\varphi\colon X \dashrightarrow Z$ is finite of degree d. Then, by Proposition 7.7, $f\varphi\colon X \dashrightarrow Z$ is a covering map of degree d if we restrict Z to a dense subset $U_\varphi \subset Z$ which contains an open subset of Z. Thus, $f\varphi$ for all $\varphi \in \mathtt{Bir}(f)$ are covering maps if we restrict Z to a dense subset

$$U_{\mathtt{Bir}} = \bigcap_{\varphi \in \mathtt{Bir}(f)} U_\varphi \subset Z$$

which contains an open subset of Z. Since φ preserves the fiber of f, then $f^{-1}(U_{\mathtt{Bir}}) = (f\varphi)^{-1}(U_{\mathtt{Bir}})$ for all $\varphi \in \mathtt{Bir}(f)$. Now, both $\mathtt{Bir}(f)$ and $\mathtt{Deck}(f_c)$ act on the dense subset $f^{-1}(U_{\mathtt{Bir}}) \subset X$ which contains an open subset of X. The restriction map

$$\begin{aligned} r_1\colon \mathtt{Bir}(f) &\to \mathtt{Bir}(f)\big|_{f^{-1}(U_{\mathtt{Bir}})} \\ \varphi &\mapsto \varphi\big|_{f^{-1}(U_{\mathtt{Bir}})} \end{aligned}$$

is injective since if two rational maps agree on a dense open subset, they agree on the intersection of their domains. Similarly, the restriction map

$$\begin{aligned} r_2\colon \mathtt{Deck}(f_c) &\to \mathtt{Deck}(f_c)\big|_{f^{-1}(U_{\mathtt{Bir}})} \\ d &\mapsto d\big|_{f^{-1}(U_{\mathtt{Bir}})} \end{aligned}$$

is injective since, by Proposition 5.20, if two deck transformations coincide on one point, they coincide on the whole $\widetilde{U}$. Since $\mathtt{Bir}(f) = \mathtt{Deck}(f_c)$, these groups have the same cardinality. Then

$$|r_1(\mathtt{Bir}(f))| = |\mathtt{Bir}(f)| = |\mathtt{Deck}(f_c)| = |r_2(\mathtt{Deck}(f_c))|.$$

Since every birational map in $r_1(\mathtt{Bir}(f))$ is a homeomorphism in the analytic topology, we conclude that

$$r_1(\mathtt{Bir}(f)) = r_2(\mathtt{Deck}(f_c)),$$

i.e. every deck transformation $d \in \mathtt{Deck}(f_c)$ in the analytic topology is indeed a birational map $\varphi \in \mathtt{Bir}(f)$.

Example 7.22. We continue with Example 7.13. We saw there that the field extension $\mathbb{C}(Z) \subset \mathbb{C}(X)$ of degree 3 is Galois with Galois group isomorphic to $\mathbb{Z}/3\mathbb{Z}$. Thus, by Proposition 7.21,

$$\mathtt{Bir}(f) = \mathrm{Gal}(\mathbb{C}(X)/\mathbb{C}(Z)) = \mathtt{Gal}(f) = \mathbb{Z}/3\mathbb{Z}.$$

It can be noticed that the generator of $\mathtt{Bir}(f)$ is the following map

$$\begin{aligned} \varphi\colon X &\dashrightarrow X \\ \begin{bmatrix} x \\ z \end{bmatrix} &\mapsto \begin{bmatrix} e^{2\pi i \frac{1}{3}} x \\ z \end{bmatrix} \end{aligned}$$

since

$$(e^{2\pi i\frac{1}{3}}x)^3 \quad z = x^3 \quad z.$$

Example 7.23. We continue with Example 7.16. We saw there that

$$\mathtt{Gal}(f) = \mathtt{Mon}_F(f_c) = \Big\langle (1\ 4), (5\ 1)(6\ 4), (2\ 4)(1\ 3) \Big\rangle.$$

Using isomorphism 7.3, we obtain

$$\mathtt{Bir}(f) = \mathrm{C}_{S_6}\Big(\Big\langle (1\ 4), (5\ 1)(6\ 4), (2\ 4)(1\ 3) \Big\rangle\Big) = \Big\langle (1\ 4)(2\ 3)(5\ 6) \Big\rangle = \mathsf{Z}/2\mathsf{Z}.$$

It can be noticed that the generator of $\mathtt{Bir}(f)$ has the form

$$\begin{array}{c} \varphi\colon X \ 99\mathsf{K}\ X \\ \begin{bmatrix} x \\ z \end{bmatrix} \ 7! \ \begin{bmatrix} \quad x \\ z \end{bmatrix} \end{array}$$

since

$$(\quad x)^6 + 2(\quad x)^4 + 3(\quad x)^2 + z = x^6 + 2x^4 + 3x^2 + z.$$

Example 7.24. Let $X = \mathbf{V}(x^2 + ax + b) \quad \mathsf{C}^3$ and $Z = \mathsf{C}^2$. Define a map

$$\begin{array}{c} f\colon X \ 99\mathsf{K}\ Z \\ \begin{bmatrix} x \\ a \\ b \end{bmatrix} \ 7! \ \begin{bmatrix} a \\ b \end{bmatrix} \end{array}$$

This map is defined everywhere on X. Is is surjective by the Fundamental Theorem of Algebra. The a□ne varieties X and Z are irreducible since $\mathbf{I}(X) = hx^2 + ax + bi$ and $\mathbf{I}(Z) = f0g$ are prime ideals. Also $\dim(X) = \dim(Z) = 2$. The Jacobian of $\mathbf{F} = \begin{bmatrix} x^2 + ax + b \end{bmatrix}$ equals

$$J(\mathbf{F}) = \begin{bmatrix} \frac{\partial \mathbf{F}}{\partial x} & \frac{\partial \mathbf{F}}{\partial a} & \frac{\partial \mathbf{F}}{\partial b} \end{bmatrix} = \begin{bmatrix} 2x + a & x & 1 \end{bmatrix}.$$

The submatrix $\begin{bmatrix} 2x + a \end{bmatrix}$ has rank less than 1 if and only if $2x + a = 0$. Thus, the image of the critical points of f are the roots of the unique generator of the elimination ideal

$$hx^2 + ax + b, 2x + ai \setminus \mathsf{C}[a, b] = ha^2 \quad 4bi.$$

Thus, $U = f(X) n f(C_f) = Z n \mathbf{V}(a^2 \quad 4b)$ and

$$f_c\colon \widetilde{U} \ ! \ U$$

is a covering map in the analytic topology. Let $u = \begin{bmatrix} 0 & 1 \end{bmatrix}^> 2\ U$. Then the covering map f_c has degree $jf_c{}^{1}(u)j = j\mathbf{V}(x^2 \quad 1)j = 2$. Hence f is a finite rational map of degree 2 and it induces an inclusion of function fields $\mathsf{C}(Z) \quad \mathsf{C}(X)$ of degree 2. We could construct a loop γ in U based at u such that it induces a transposition of the elements in the fiber $F = f_c{}^{1}(u)$. In other words, it can be shown that $\mathtt{Gal}(f) = \mathtt{Mon}_F(f_c) = \mathsf{Z}/2\mathsf{Z}$.

However, let's use a little bit different strategy for determining $\mathtt{Gal}(f)$. From Galois theory it follows that every field extension of degree 2 is Galois, i.e. the field extension $\mathbb{C}(Z) \subset \mathbb{C}(X)$ is Galois. Thus,

$$\mathtt{Mon}_F(f_c) = \mathtt{Gal}(f) = \mathrm{Gal}(\mathbb{C}(X)/\mathbb{C}(Z)).$$

From Galois theory it also follows that $F \subset L$ is a Galois extension if and only if $|\mathrm{Gal}(L/F)| = [L : F]$ (as was shown in Example 6.29). Thus,

$$|\mathrm{Gal}(\mathbb{C}(X)/\mathbb{C}(Z))| = [\mathbb{C}(X) : \mathbb{C}(Z)] = 2 \Rightarrow \mathrm{Gal}(\mathbb{C}(X)/\mathbb{C}(Z)) = \mathbb{Z}/2\mathbb{Z}.$$

By Proposition 7.21, we have

$$\mathtt{Bir}(f) = \mathrm{Gal}(\mathbb{C}(X)/\mathbb{C}(Z)) = \mathbb{Z}/2\mathbb{Z}.$$

It can be noticed that the generator of $\mathtt{Bir}(f)$ has the following form

$$\begin{aligned} \varphi\colon X &\dashrightarrow X \\ \begin{bmatrix} x \\ a \\ b \end{bmatrix} &\mapsto \begin{bmatrix} x & a \\ & a \\ & b \end{bmatrix} \end{aligned}$$

since

$$(x-a)^2 + a(x-a) + b = x^2 + 2ax + a^2 - ax - a^2 + b = x^2 + ax + b.$$

Despite the fact that the roots $x_{1,2} = \frac{a \pm \sqrt{a^2 - 4b}}{2}$ of $x^2 + ax + b$ are not rational functions, the function $x \mapsto x - a$ which interchanges them is rational.

In Examples 7.22, 7.23, 7.24 we have shown how we can reveal the structure of $\mathtt{Bir}(f)$. However, we didn't explain how the generators of $\mathtt{Bir}(f)$ can be found. We will give an ideal of how it can be done in Chapter 9.

8 Imprimitivity of Galois Group

In Chapter 7 we saw how the Galois group of a branched cover can be computed and how the symmetries can be revealed from it. In this chapter we will show how to use these symmetries for passing from the original branched cover to a branched cover with smaller degree. In Chapter 9 we will explain how this relates to polynomial system simplification.

If a branched cover f has symmetries, then we can factorize it and obtain two branched covers of strictly smaller degree. In other words, f is decomposable (see Definition 8.7). However, it may happen that f has no symmetries (i.e. $\mathtt{Bir}(f)$ is trivial), but it is still decomposable. Thus, the existence of symmetries is not a necessary condition for f to be decomposable. In this chapter we show that there exists a necessary and su□cient condition for f to be decomposable. This condition is associated with one of the properties of the Galois/monodromy group of f, namely imprimitivity.

8.1 Imprimitive permutation groups

Let G S_d be a permutation group.

Definition 8.1. A block of G is a subset B $f1,\dots,dg$ such that for every g 2 G, either $gB = B$ or $gB \setminus B = ?$.

The subsets ?, $f1,\dots,dg$, and every singleton are blocks of every permutation group. These blocks are called trivial.

Definition 8.2. A transitive permutation group G S_d is called imprimitive if there exists a nontrivial block of G. Otherwise, it is primitive.

Example 8.3. Let $G = h(1\ 2\ 3\ 4)i$ S_4 be a subgroup of order 4. It is obvious that G is transitive. Then a subset $B_1 = f1,3g$ $f1,\dots,4g$ is a nontrivial block of G. Thus, G is imprimitive. Notice that $B_2 = f2,4g$ $f1,\dots,4g$ is also a nontrivial block of G.

It can be noticed that B_1 and B_2 from Example 8.3 form a partition of the set $f1,\dots,4g$ into blocks of size 2. We give the following proposition which shows that every imprimitive group forms such a partition.

Proposition 8.4. *If G is imprimitive, then there is a decomposition $B = fB_ig_{i=1}^m$ of the set $f1,\dots,dg$ into disjoint nontrivial blocks of G of size k, i.e. $d = mk$.*

Proof. Let B be a nontrivial block of G. Then define

$$B = \{gB \mid g \in G\}.$$

Take $g_1, g_2 \in G$ and let $g_1B, g_2B \in B$. We need to prove that

$$g_1B = g_2B \text{ or } g_1B \cap g_2B = \emptyset \tag{8.1}$$

Because B is a block of G and $g_2^{-1}g_1 \in G$, then

$$g_2^{-1}g_1B = B \text{ or } g_2^{-1}g_1B \cap B = \emptyset \tag{8.2}$$

However, (8.2) is equivalent to (8.1) since

$$g_1B = g_2B \quad (\Leftrightarrow) \quad g_2^{-1}g_1B = g_2^{-1}g_2B = B,$$

$$g_1B \cap g_2B = \emptyset \quad (\Leftrightarrow) \quad g_2^{-1}g_1B \cap g_2^{-1}g_2B = g_2^{-1}g_1B \cap B = \emptyset.$$

So, we have shown that B is a partition of $\{1, \ldots, d\}$ into disjoint subsets. These subsets have equal size since $|gB| = |B|$ (multiplication by g is injective map). Now, to show that the elements of B are blocks of G, take $g_1B \in B$ and $g_2 \in G$. We need to prove that

$$g_2g_1B = g_1B \text{ or } g_2g_1B \cap g_1B = \emptyset.$$

The above statement is equivalent to

$$g_1^{-1}g_2g_1B = B \text{ or } g_1^{-1}g_2g_1B \cap B = \emptyset,$$

which is true since B is a block of G. □

We state another useful fact about imprimitive groups.

Proposition 8.5. *A permutation group $G \leq S_d$ is imprimitive if and only if there is a proper subgroup of G which strictly contains $\mathrm{Stab}_G(1)$ (i.e. $\mathrm{Stab}_G(1)$ is not a maximal subgroup of G).*

Proof. [3, Section 1.1, the last paragraph]. □

In fact, all the stabilizers $\mathrm{Stab}_G(x)$ for $x \in \{1, \ldots, d\}$ of a transitive permutation group G are conjugate, i.e.

$$\mathrm{Stab}_G(x_2) = g\,\mathrm{Stab}_G(x_1)\,g^{-1},$$

for $g(x_1) = x_2$. We know that such an element g exists since G is transitive. If $\mathrm{Stab}_G(x_1)$ is not a maximal subgroup of G, then there is a proper subgroup H_1 of G which strictly contains $\mathrm{Stab}_G(x_1)$. But then

$$H_2 = gH_1g^{-1}$$

is a proper subgroup of G which strictly contains $\mathrm{Stab}_G(x_2)$. In other words, $\mathrm{Stab}_G(x_2)$ is not a maximal subgroup of G. Thus, Proposition 8.5 can be restated as follows.

Proposition 8.6. *A permutation group $G \leq S_d$ is imprimitive if and only if $\mathrm{Stab}_G(x)$ is not a maximal subgroup of G for all $x \in \{1, \ldots, d\}$.*

8.2 Decomposable branched covers

Definition 8.7. A branched cover $f\colon X \dashrightarrow Z$ of irreducible a□ne varieties is called decomposable if there exists an irreducible a□ne variety Y such that f factors

$$X \overset{f_1}{\dashrightarrow} Y \overset{f_2}{\dashrightarrow} Z \tag{8.3}$$

where f_1 and f_2 are finite rational maps with degrees $d_1, d_2 > 1$, respectively, such that $f = f_2 f_1$.

Example 8.8. Let f be the branched cover from Example 7.16. Then f factors as

$$X \overset{f_1}{\dashrightarrow} Y \overset{f_2}{\dashrightarrow} Z$$

$$(x, z) \longmapsto (x^2, z) \longmapsto z$$

for $Y = \mathbf{V}(y^3 + 2y^2 + 3y + z)$ C^2.

Remark 8.9. From Corollary 6.40 we know that there is a one-to-one correspondence between intermediate fields $\mathsf{C}(Z)$ K $\mathsf{C}(X)$ and intermediate groups $\mathrm{Gal}(N_f/\mathsf{C}(X))$ H $\mathtt{Gal}(f)$. Since $\mathsf{C}(Z)$ and $\mathsf{C}(X)$, by (6.2), correspond to $\mathtt{Gal}(f)$ and $\mathrm{Gal}(N_f/\mathsf{C}(X))$, respectively, there is a one-to-one correspondence between intermediate fields

$$\mathsf{C}(Z) \subsetneq K \subsetneq \mathsf{C}(X)$$

and intermediate groups

$$\mathrm{Gal}(N_f/\mathsf{C}(X)) \subsetneq H \subsetneq \mathtt{Gal}(f).$$

By the first paragraph from Section 7.2, $\mathsf{C}(X) = \mathsf{C}(Z)(\alpha_1)$, and thus $\mathrm{Gal}(N_f/\mathsf{C}(X))$ consists of elements of $\mathtt{Gal}(f)$ which fix α_1. In other words,

$$\mathrm{Gal}(N_f/\mathsf{C}(X)) \overset{\chi}{=} \mathrm{Stab}_{\chi(\mathtt{Gal}(f))}(\alpha_1).$$

Hence there is a one-to-one correspondence between intermediate fields

$$\mathsf{C}(Z) \subsetneq K \subsetneq \mathsf{C}(X)$$

and intermediate groups

$$\mathrm{Stab}_{\chi(\mathtt{Gal}(f))}(\alpha_1) \subsetneq H \subsetneq \chi(\mathtt{Gal}(f)).$$

We give the following proposition which states a necessary and su□cient condition for a branched cover to be decomposable. In the statement of this proposition we identify $\mathtt{Gal}(f)$ with $\chi(\mathtt{Gal}(f))$.

Proposition 8.10. *A branched cover $f\colon X \dashrightarrow Z$ is decomposable if and only if its Galois group $\mathtt{Gal}(f)$ is imprimitive.*

Proof. By Proposition 8.5, imprimitivity of $\chi(\mathtt{Gal}(f))$ is equivalent to the existence of an intermediate group

$$\mathrm{Stab}_{\chi(\mathtt{Gal}(f))}(\alpha_1) \subsetneq H \subsetneq \chi(\mathtt{Gal}(f)).$$

This, by Remark 8.9, is equivalent to the existence of an intermediate field

$$\mathsf{C}(Z) \subsetneq K \subsetneq \mathsf{C}(X).$$

Since $\mathsf{C}(X)$ is finitely generated over C (by coordinate functions $x_1, \ldots, x_n \in \mathsf{C}(X)$), then K is finitely generated over C [4, Theorem 1.1]. Then it follows that $K = \mathsf{C}(Y)$ for some irreducible a□ne variety [20, Lemma 1.3]. By Propositions 4.43 and 7.10, this is equivalent to the existence of finite rational maps f_1, f_2 with degrees $d_1 = [\mathsf{C}(X) : K] > 1$ and $d_2 = [K : \mathsf{C}(Z)] > 1$, respectively, such that f factors

$$X \overset{f_1}{\dashrightarrow} Y \overset{f_2}{\dashrightarrow} Z$$

By Corollary 6.40, we have

$$[\mathsf{C}(X) : K] = [\mathrm{Gal}(N_f/K) : \mathrm{Gal}(N_f/\mathsf{C}(X))] = [H : \mathrm{Stab}_{\chi(\mathtt{Gal}(f))}(\alpha_1)],$$

$$[K : \mathsf{C}(Z)] = [\mathrm{Gal}(N_f/\mathsf{C}(Z)) : \mathrm{Gal}(N_f/K)] = [\mathtt{Gal}(f) : H].$$

By Proposition 7.10,

$$d_1 = [\mathsf{C}(X) : K], \quad d_2 = [K : \mathsf{C}(Z)].$$

□

Example 8.11. We continue with Example 7.23. We have $X = \mathbf{V}(x^6 + 2x^4 + 3x^2 + z) \subset \mathsf{C}^2$, $Z = \mathsf{C}$ and

$$f \colon X \dashrightarrow Z$$
$$\begin{bmatrix} x \\ z \end{bmatrix} \mapsto z$$

We know that

$$\mathtt{Gal}(f) = \mathtt{Mon}_F(f_c) = \Big\langle (1\ 4), (5\ 1)(6\ 4), (2\ 4)(1\ 3) \Big\rangle.$$

It can be seen that $\{2, 3\}$ is a nontrivial block of $\mathtt{Gal}(f)$ and, thus, by Definition 8.2, $\mathtt{Gal}(f)$ is imprimitive. By the proof of Proposition 8.4, the partition of the set $\{1, \ldots, 6\}$ into blocks of $\mathtt{Gal}(f)$ is

$$B = \Big\{ \{1, 4\}, \{2, 3\}, \{5, 6\} \Big\}.$$

Let $\mathrm{Stab}_{\mathtt{Gal}(f)}(1)$ be the stabilizer of $1 \in \{1, \ldots, 6\}$ by $\mathtt{Gal}(f)$. By Remark 8.9, there is a one-to-one correspondence between intermediate subgroups $\mathrm{Stab}_{\mathtt{Gal}(f)}(1) \subsetneq H \subsetneq \mathtt{Gal}(f)$ and intermediate fields $\mathsf{C}(Z) \subsetneq K \subsetneq \mathsf{C}(X)$. In GAP we compute all the intermediate subgroups H using the following code

```
Gal := Group((1,4),(5,1)(6,4),(2,4)(1,3));
Stab := Stabilizer(Gal,1);
IntermediateSubgroups(Gal,Stab);
```

It turns out that there is only one such H, namely

$$H = h(2\ 3), (2\ 6)(3\ 5), (1\ 4)(5\ 6)i.$$

The stabilizer $\mathrm{Stab}_{\mathtt{Gal}(f)}(1)$ equals

$$\mathrm{Stab}_{\mathtt{Gal}(f)}(1) = h(2\ 3), (2\ 6)(3\ 5)i.$$

Thus, there is only one intermediate field $\mathsf{C}(Z)$ (K ($\mathsf{C}(X)$. By Proposition 8.10, f is decomposable. We can notice that f factors as

$$X \xrightarrow{\ f_1\ } Y \xrightarrow{\ f_2\ } Z$$

$$(x, z) \longmapsto (x^2, z) \longmapsto z$$

for $Y = \mathbf{V}(y^3 + 2y^2 + 3y + z)$ C^2 with degrees

$$d_1 = [H : \mathrm{Stab}_{\mathtt{Gal}(f)}(1)] = \frac{jHj}{j\mathrm{Stab}_{\mathtt{Gal}(f)}(1)} = \frac{16}{8} = 2,$$

$$d_2 = [\mathtt{Gal}(f) : H] = \frac{j\mathtt{Gal}(f)j}{jHj} = \frac{48}{16} = 3,$$

respectively. Also $\mathsf{C}(Y) = K$. Of course, the question is how do we notice that for more complicated branched covers? For this we can use an argument given in [3, p. 4]. Our aim is to find the generators of the intermediate field K over C (it is finitely generated over C since $\mathsf{C}(X)$ is). Then the map f_1 will be given by these generators. Recall from Example 7.23 that $\mathtt{Bir}(f)$ is generated by the element

$$\begin{aligned} \varphi\colon X &\dashrightarrow X \\ \begin{bmatrix} x \\ z \end{bmatrix} &\mapsto \begin{bmatrix} x \\ z \end{bmatrix} \end{aligned}$$

By Proposition 7.21, $\mathtt{Bir}(f)$, identified with $\beta(\mathtt{Bir}(f))$, acts on $\mathsf{C}(X)$ by automorphisms fixing $\mathsf{C}(Z)$ point-wise. We can look at the fixed subfield of this group:

$$\mathsf{C}(X)_{\mathtt{Bir}(f)} = f g \ 2\ \mathsf{C}(X)\ j\ \psi(g) = g \quad \text{for all } \psi\ 2\ \mathtt{Bir}(f)g.$$

Since the order of $\mathtt{Bir}(f)$ equals 2, then [5, Chapter 7, Theorem 7.5.3] shows that

$$[\mathsf{C}(X) : \mathsf{C}(X)_{\mathtt{Bir}(f)}] = j\mathtt{Bir}(f)j = 2.$$

In other words, the field extension $\mathsf{C}(X)_{\mathtt{Bir}(f)}$ $\mathsf{C}(X)$ is finite of degree 2. Since $[\mathsf{C}(X) : \mathsf{C}(Z)] = 6$, then $\mathsf{C}(Z)$ ($\mathsf{C}(X)_{\mathtt{Bir}(f)}$ ($\mathsf{C}(X)$, and thus

$$\mathsf{C}(X)_{\mathtt{Bir}(f)} = K.$$

Since φ is actually a linear map, then we can use invariant theory (see [6, Chapter 7]) to compute the generators of $\mathsf{C}(X)_{\mathtt{Bir}(f)}$ over C. These are $x^2, z\ 2\ \mathsf{C}(X)_{\mathtt{Bir}(f)}$. Thus, f_1 is

given by x^2 and z, as was mentioned before. To compute the equations defining Y can be computed using polynomial implicitization (see [6, Chapter 3, §3]). We consider the ideal

$$I = \langle x^6 + 2x^4 + 3x^2 + z, y_1 \quad x^2, y_2 \quad z \rangle.$$

To find the equations defining Y we need to eliminate x and z from I, i.e.

$$J = I \cap \mathbb{C}[y_1, y_2] = \langle y_1^3 + 2y_1^2 + 3y_1 + y_2 \rangle.$$

We can go back to look at the structure of $\mathtt{Gal}(f)$. It has 48 elements. Indeed, it is permutation isomorphic to a wreath product

$$G_f = S_2 \wr S_3$$

constructed in Example 3.40.

We give a general fact about imprimitive groups.

Proposition 8.12. *Let G be an imprimitive group that admits a complete block system B consisting of m blocks of size k. Then G is permutation isomorphic to a subgroup of $S_k \wr S_m$.*

Proof. [8, Theorem 5.4]. □

9 Applications to Solving Point-Line Minimal Problems in Computer Vision

In [11], Table 1, there is a complete list of point-line minimal problems in complete multiview visibility. In this chapter we're going to explore three of them: 5000_2 with 20 solutions (also known as 5-point relative pose problem), 3100_0 with 64 solutions and 3010_0 with 216 solutions. We will explain theoretical details of the known reduction of the 5-point problem using the essential matrix. It is also explained why the formulation with essential matrix cannot be simplified more. It was discovered [9] that the problem 3100_0 can be reduced to a simpler problem with 16 solutions. Concerning the problem 3010_0, it was discovered [23] that it cannot be reduced to a simpler problem with less number of solutions.

In Chapters 4 and 7 we considered a□ne varieties as the main object of the study since they are in some sense simple to imagine and understand. However, when it comes to solving minimal problems in computer vision, we have to switch the language of a□ne varieties to something more general: the language of quasiprojective varieties [24, Section 4]. These are the intersections of a Zariski-open and a Zariski-closed subset inside some projective space. Fortunately, everything described in Chapters 4 and 7 can be restated for quasiprojective varieties too.

9.1 Point-line minimal problem 5000_2

Point-line minimal problem 5000_2 (or 5-point relative pose problem) is the problem of estimating the camera relative pose from 5 point correspondences in 2 views. We can formulate this problem as follows:

$$\begin{aligned} \mathsf{R}^\top\mathsf{R} = \mathsf{I}, \quad \det(\mathsf{R}) = 1, \\ \beta_i \begin{bmatrix} \mathbf{y}_i \\ 1 \end{bmatrix} = \mathsf{R}\alpha_i \begin{bmatrix} \mathbf{x}_i \\ 1 \end{bmatrix} + \mathbf{t}, \quad i = 1, \dots, 5. \end{aligned} \tag{9.1}$$

Here $(\mathsf{R}, \mathbf{t})$ is the relative pose between two cameras (See Figure 9.1). The tuples $(\mathbf{x}_i, \mathbf{y}_i)$ are the 5 image correspondences in the 1st and the 2nd camera, respectively. Depths α_i and β_i are the quotients $\frac{k\mathbf{X}_i k}{k\left[\begin{smallmatrix}\mathbf{x}_i \\ 1\end{smallmatrix}\right]k}$ and $\frac{k\mathsf{R}\mathbf{X}_i + \mathbf{t}k}{k\left[\begin{smallmatrix}\mathbf{y}_i \\ 1\end{smallmatrix}\right]k}$ for a 3D point $\mathbf{X}_i$, respectively. We can understand the equations (9.1) in two di□erent ways. The first way is that we pretend that $\mathbf{x}_i$ and $\mathbf{y}_i$ are vectors of real numbers (they define the certain instance of the 5pt problem) and thus

(9.1) defines an ideal in the polynomial ring $\mathbb{C}[\mathsf{R}, \mathbf{t}, \alpha_1, \ldots, \beta_5]$. The second way is that we understand $\mathbf{x}_i$ and $\mathbf{y}_i$ as variables and hence (9.1) defines an ideal

$$I \quad \mathbb{C}[\mathsf{R}, \mathbf{t}, \alpha_1, \ldots, \beta_5, \mathbf{x}_1, \ldots, \mathbf{y}_5].$$

Since the equations (9.1) are $(\mathbf{t}, \alpha_1, \ldots, \beta_5)$-homogeneous, this ideal defines a quasiprojective variety over the complex numbers

$$X = \mathbf{V}(I) \quad \mathrm{SO}(3, \mathbb{C}) \quad \mathbb{P}(\mathbb{C}^{13}) \quad \mathbb{C}^{20}.$$

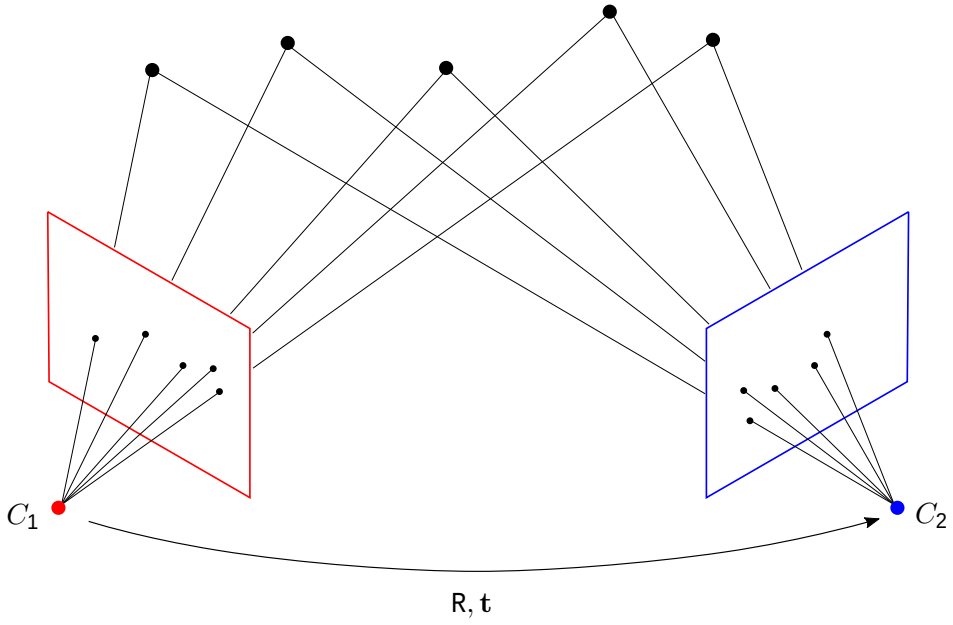


Figure 9.1

Let $Z = \mathbb{C}^{20}$. We can define a map

$$f : X \;!\; Z$$
$$\begin{bmatrix} \mathsf{R} \\ \mathbf{t} \\ \alpha_1 \\ \vdots \\ \beta_5 \\ \mathbf{x}_1 \\ \vdots \\ \mathbf{y}_5 \end{bmatrix} \;\; 7! \;\; \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{y}_5 \end{bmatrix}$$

Since f is just the projection (to the image measurements), it is obviously a rational map. Let $\widetilde{X}$ be an a□ne chart of X where we fix the last coordinate t_3 of $\mathbf{t}$ to 1. Then Also let $\widetilde{f}$ be the restricted map f between a□ne varieties $\widetilde{X}$ and Z. It can be verified computationally that

1. $\dim \widetilde{X} = \dim Z = 20$ (see [6, Chapter 9]).
2. $\widetilde{f}$ is dominant (according to the Elimination Theorem [6, Chapter 3], it is su□cient to verify that $(I + ht_3 \quad 1i) \setminus \mathsf{C}[\mathbf{x}_1, \ldots, \mathbf{y}_5] = f0g$).

It is obvious that Z is irreducible in the Zariski topology.

Remark 9.1. Also, after conversation with Tim Du□, we conclude that $\widetilde{X}$ (and hence X) is irreducible too. This is because equations (9.1) are linear in $\mathbf{x}_i$ and $\mathbf{y}_i$ and the monodromy group of $\widetilde{f}$ acts transitively on the generic fiber (for the explanation see [10, Chapter 2]).

Thus, by Proposition 7.7, there is a proper closed subset $C_{\widetilde{f}} \quad \widetilde{X}$ such that for $U \stackrel{\text{def}}{=} \widetilde{f}(\widetilde{X}) n \widetilde{f}(C_{\widetilde{f}})$ the map

$$\widetilde{f}\Big|_{\widetilde{f}^{\ 1}(U)} : \widetilde{f}^{\ 1}(U) \ ! \ U$$

is a covering map (Definition 5.10) in the analytic topology (Section 7.1).

We would like to show that this covering has degree 20. For this, according to Remark 5.11, it is su□cient to take one point $z_0 \ 2 \ U$ and show that $\widetilde{f}^{\ 1}(z_0)$ consists of 20 distinct points, since, by Proposition 7.7, U is connected in the analytic topology. However, it is hard to sample from U. Usually, in practice this is done by taking a generic (randomly chosen) point $z_0 \ 2 \ Z$ and computing the cardinality $j\widetilde{f}^{\ 1}(z_0)j$ using Groebner bases techniques. Since $U \quad Z$ contains some open subset of Z and Z is irreducible, then U is "almost the whole Z". In other words, there is almost 100% probability that $z_0 \ 2 \ U$. However, it may happen that $z_0 \ \not{2} \ U$, i.e. the degree of the covering map may be di□erent from $j\widetilde{f}^{\ 1}(z_0)j$. Thus, after computing $\widetilde{f}^{\ 1}(z_0)$ we have to check that $z_0 \ 2 \ U$. So, we take a generic point $z_0 \ 2 \ Z$ and verify that $j\widetilde{f}^{\ 1}(z_0)j = 20$ using Groebner bases techniques. Let $\mathbf{F}$ denotes the column vector of 23 polynomials from (9.1) including the polynomial $t_3 \quad 1$. To show that $z_0 \ 2 \ U$ we need to check that for every $x \ 2 \ \widetilde{f}^{\ 1}(z_0)$ the matrix

$$J_{\mathsf{R},\mathbf{t},\alpha_1,\ldots,\beta_5}(\mathbf{F})_x \stackrel{\text{def}}{=} \left[\frac{\partial \mathbf{F}}{\partial r_{11}} \cdots \frac{\partial \mathbf{F}}{\partial r_{33}} \frac{\partial \mathbf{F}}{\partial t_1} \cdots \frac{\partial \mathbf{F}}{\partial t_3} \frac{\partial \mathbf{F}}{\partial \alpha_1} \cdots \frac{\partial \mathbf{F}}{\partial \beta_5}\right]_x \ 2 \ \mathsf{C}^{23 \quad 22}$$

has rank 22.

By Proposition 7.7, U contains an open subset V of Z. Since Z is irreducible, V is dense in Z. Since for every $z_0 \ 2 \ U$ the fiber $\widetilde{f}^{\ 1}(z_0)$ consists of 20 distinct points, then, by Definition 7.9, $\widetilde{f}$ is a finite map of degree 20. Thus, by Proposition 7.10, $\widetilde{f}$ induces an inclusion of function fields $\widetilde{f} : \mathsf{C}(Z) \ ! \ \mathsf{C}(\widetilde{X})$ of degree 20. Since $\widetilde{X}$ is dense in X, then $\mathsf{C}(X) = \mathsf{C}(\widetilde{X})$ and thus f induces an inclusion of function fields $f : \mathsf{C}(Z) \ ! \ \mathsf{C}(X)$ of the same degree 20. Again, by Proposition 7.10 (just restated for quasiprojective varieties), f is finite of degree 20.

We can now compute the Galois/monodromy group $\mathtt{Gal}(f)$ of f as was described in Section 7.2. Since $\mathsf{C}(X) = \mathsf{C}(\widetilde{X})$, then $\mathtt{Gal}(f) = \mathtt{Gal}(\widetilde{f})$ (this is because isomorphic field

extensions have isomorphic Galois closures and thus the Galois groups of these closures are isomorphic [5, Proposition 6.1.11]). So, by computing $\mathtt{Gal}(\widetilde{f})$ we get $\mathtt{Gal}(f)$. The Galois group $\mathtt{Gal}(\widetilde{f})$ has been computed by [23] using [16]:

$$\mathtt{Gal}(f) = \mathtt{Gal}(\widetilde{f}) = \Big\langle (2\ 9\ 6\ 18\ 8\ 16\ 15\ 10)(3\ 4\ 7\ 17\ 19\ 11\ 14\ 5),$$

$$(1\ 13\ 17\ 11)(2\ 20\ 12\ 15)(3\ 19\ 9\ 18\ 10\ 14)(4\ 16\ 8\ 7)\Big\rangle.$$

Using GAP we can compute a decomposition of the set $\{1,\dots,20\}$ into blocks of $\mathtt{Gal}(f)$ (see Definition 8.1 and Proposition 8.4). The GAP code which does that is

```
Galois_Group := Group((2,9,6,18,8,16,15,10)(3,4,7,17,19,11,14,5),
(1,13,17,11)(2,20,12,15)(3,19,9,18,10,14)(4,16,8,7));
Blocks(Galois_Group,[1..20]);
```

The result of the last command is the decomposition of $\{1,\dots,20\}$ into 10 blocks of $\mathtt{Gal}(f)$ of size 2:

$$B = \Big\{\{1,20\},\{2,11\},\{3,18\},\{4,8\},\{5,6\},\{7,16\},\{9,14\},\{10,19\},\{12,13\},\{15,17\}\Big\}. \tag{9.2}$$

This reflects the fact that the formulation (9.1) of the 5pt problem can be reduced to a simpler problem (formulation using the essential matrix) with 10 solutions. By Definition 8.2, $\mathtt{Gal}(f)$ is imprimitive. By Proposition 8.10, f is decomposable, which means that there is a quasiprojective variety Y such that f factors

$$\begin{array}{ccccc} X & \overset{f_1}{\dashrightarrow} & Y & \overset{f_2}{\dashrightarrow} & Z \\ \mathsf{C}(X) & \overset{f_1}{\longleftarrow} & \mathsf{C}(Y) & \overset{f_2}{\longleftarrow} & \mathsf{C}(Z) \end{array} \tag{9.3}$$

where f_1 and f_2 are finite rational maps with degrees $d_1 = 2$ and $d_2 = \frac{20}{2} = 10$, respectively, such that $f = f_2 \quad f_1$. It is important to say that f_1,f_2 and Y are not unique. We can transform Y by any birational map g:

$$\begin{array}{ccccc} X & \overset{f_1}{\dashrightarrow} & Y & \overset{f_2}{\dashrightarrow} & Z \\ & & \downarrow g & & \\ & & \widetilde{Y} & & \end{array}$$

and thus obtain

$$X \overset{g\ f_1}{\dashrightarrow} \widetilde{Y} \overset{f_2\ g\ ^1}{\dashrightarrow} Z$$

Thus, there is a question how to find an optimal map f_1 so that the equations defining Y are simple, but we don't solve this problem here. It is well-known that one of the choices of

(actually, polynomial) f_1 and f_2 is the following one

$$X \xrightarrow{f_1} Y \xrightarrow{f_2} Z$$

$$\begin{bmatrix} \mathsf{R} \\ \mathbf{t} \\ \alpha_1 \\ \vdots \\ \beta_5 \\ \mathbf{x}_1 \\ \vdots \\ \mathbf{y}_5 \end{bmatrix} \longmapsto \begin{bmatrix} [\mathbf{t}] & \mathsf{R} \\ \mathbf{x}_1 & \\ \vdots & \\ \mathbf{y}_5 & \end{bmatrix} \longmapsto \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{y}_5 \end{bmatrix} \tag{9.4}$$

where $[\mathbf{t}]$ denotes the skew-symmetric matrix

$$[\mathbf{t}] = \begin{bmatrix} 0 & t_3 & t_2 \\ t_3 & 0 & t_1 \\ t_2 & t_1 & 0 \end{bmatrix}.$$

It is obvious that $f = f_2 \quad f_1$. It is well-known that $\deg(f_1) = 2$. Thus, $\deg(f_2) = \frac{20}{2} = 10$.

What remains is to find the equations defining Y. It can be done using rational (or polynomial, for our choice of f_1) implicitization (see [6, Chapter 3, §3] for a□ne varieties). Let

$$\mathsf{E} = \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix}$$

be the matrix of 9 indeterminates and

$$\widetilde{I} = \left\langle \mathsf{R}^{>}\mathsf{R} \quad \mathtt{I}, \det(\mathsf{R}) \quad 1, \beta_i \begin{bmatrix} \mathbf{y}_i \\ 1 \end{bmatrix} \quad \mathsf{R}\alpha_i \begin{bmatrix} \mathbf{x}_i \\ 1 \end{bmatrix} \quad \mathbf{t}, \mathsf{E} \quad [\mathbf{t}] \quad \mathsf{R} \right\rangle, \quad i = 1, \ldots, 5.$$

The ideal defining Y can be computed as follows:

$$J = \widetilde{I} \setminus \mathsf{C}[\mathsf{E}, \mathbf{x}_1, \ldots, \mathbf{y}_5] \tag{9.5}$$

This elimination ideal equals

$$J = \left\langle 2\mathsf{E}\mathsf{E}^{>}\mathsf{E} \quad \operatorname{trace}(\mathsf{E}^{>}\mathsf{E})\mathsf{E}, \ \det(\mathsf{E}), \ \begin{bmatrix} \mathbf{y}_i^{>} & 1 \end{bmatrix} \mathsf{E} \begin{bmatrix} \mathbf{x}_i \\ 1 \end{bmatrix} \right\rangle, \quad i = 1, \ldots, 5. \tag{9.6}$$

Ideal 9.6 defines a quasiprojective variety

$$Y = \mathbf{V}(J) \quad \mathsf{P}(\mathsf{C}^9) \quad \mathsf{C}^{20} \tag{9.7}$$

since the equations (9.6) are E-homogeneous. We conclude with the following observation: $f^{-1}(z) = f_1^{-1}(f_2^{-1}(z))$ which means that instead of solving (9.1) with 20 solutions, we can solve (9.6) with 10 solutions, take $q \quad 10$ real solutions and for each of them find the fiber

under f_1. The latter will be more e□cient than finding all the 20 solutions directly from (9.1).

We can also compute the Galois/monodromy group of f_2. It turns out that $\mathtt{Gal}(f_2) = S_{10}$. By Definition 8.2, S_{10} is primitive since it doesn't have a nontrivial block. It means, by Proposition 8.10, that f_2 is indecomposable, i.e. there doesn't exist a quasiprojective variety Y^0 such that f_2 factors

$$Y \xrightarrow{f_3} Y^0 \xrightarrow{f_4} Z$$

as a composition of finite rational maps f_3, f_4 of degrees $d_3, d_4 > 1$, respectively, such that $f_2 = f_4 \quad f_3$.

Remark 9.2. It is also known that S_{10} is not 9-solvable. From Galois theory it follows that in the Galois closure N_{f_2} of $\mathsf{C}(Y)/\mathsf{C}(Z)$ there is no tower of fields

$$\mathsf{C}(Z) = F_0 \quad F_1 \qquad\qquad F_{r\ 1} \quad F_r$$

with $N_{f_2} \quad F_r$ such that $F_{i+1} = F_i(a_i)$, where a_i is a solution of an algebraic equation of degree at most 9 over F_i, or satisfies $a_i^m = b_i$ for some natural number m and b_i *2* F_i. In particular, this means that the rational functions $\frac{e_{12}}{e_{11}}, \ldots, \frac{e_{33}}{e_{11}}$ *2* $\mathsf{C}(Y)$ cannot be expressed in terms of $\mathbf{x}_1, \ldots, \mathbf{y}_5$ *2* $\mathsf{C}(Z)$ using the operations of addition, subtraction, multiplication, division, solving univariate polynomials of degree at most 9 and extracting roots. In other words, the formulation using the essential matrix $\mathtt{E}$ cannot be simplified more. Similar verification was made in [21]. There, basically, the Galois group G was computed symbolically over $\mathbb{Q}$ for a univariate polynomial of degree 10 from the ideal (9.6). We believe that there is a strong connection between the properties of G and $\mathtt{Gal}(f)$ in general, but the exact relationship between them is not clear to us right now.

There is still one question which remains unsolved: how do we find f_1 in general? This is where Section 7.3 can help: it may happen that a decomposable branched cover f has symmetries (see Definition 7.19). From (7.3) we know that $\mathtt{Bir}(f)$ is isomorphic to the centralizer of $\mathtt{Gal}(f)$ in S_{20}. We use GAP to compute it. The GAP command which does that is

```
Bir := Centralizer(SymmetricGroup(20), Galois_Group);
```

We obtain

$$\mathtt{Bir}(f) = \mathrm{C}_{S_{20}}(\mathtt{Gal}(f)) = \Big\langle (1\ 20)(2\ 11)(3\ 18)(4\ 8)(5\ 6)(7\ 16)(9\ 14)(10\ 19)(12\ 13)(15\ 17) \Big\rangle.$$

Thus, there is a birational map $\varphi\colon X$ 99K X of order 2 such that $f \quad \varphi = f$.

Remark 9.3. We would like to point out that it is not a coincidence that $\mathtt{Bir}(f)$ is nontrivial in our case of f: it is always nontrivial if there exists a partition of the solution set into blocks of $\mathtt{Gal}(f)$ of size 2 (like in (9.2)). In that case the nontriviality of $\mathtt{Bir}(f)$ follows from the fact that every field extension of degree 2 is Galois.

If we collect the monodromy groups for di□erent fibers we can find out how $\mathtt{Bir}(f)$ acts on each of these fibers. Making an assumption on maximal degrees of the numerators and denominators of φ, the coe□cients of φ can be computed using linear algebra. We haven't

yet tested it for this case of f, but we believe that using this method it is possible to obtain the generator of $\mathtt{Bir}(f)$:

$$X \dashrightarrow^{\varphi} X$$

$$\begin{bmatrix} R \\ t \\ \alpha_1 \\ \beta_1 \\ \vdots \\ \alpha_5 \\ \beta_5 \\ x_1 \\ \vdots \\ y_5 \end{bmatrix} \longmapsto \begin{bmatrix} \left(2\frac{tt^{>}}{t^{>}t} \quad I\right) R \\ t \\ \frac{\alpha_1 k t k^2}{k R^{>} t + \alpha_1 \left[\begin{smallmatrix} x_1 \\ 1 \end{smallmatrix}\right] k^2 \quad k \alpha_1 \left[\begin{smallmatrix} x_1 \\ 1 \end{smallmatrix}\right] k^2} \\ \frac{\beta_1 k t k^2}{k R^{>} t + \alpha_1 \left[\begin{smallmatrix} x_1 \\ 1 \end{smallmatrix}\right] k^2 \quad k \alpha_1 \left[\begin{smallmatrix} x_1 \\ 1 \end{smallmatrix}\right] k^2} \\ \vdots \\ \frac{\alpha_5 k t k^2}{k R^{>} t + \alpha_5 \left[\begin{smallmatrix} x_5 \\ 1 \end{smallmatrix}\right] k^2 \quad k \alpha_5 \left[\begin{smallmatrix} x_5 \\ 1 \end{smallmatrix}\right] k^2} \\ \frac{\beta_5 k t k^2}{k R^{>} t + \alpha_5 \left[\begin{smallmatrix} x_5 \\ 1 \end{smallmatrix}\right] k^2 \quad k \alpha_5 \left[\begin{smallmatrix} x_5 \\ 1 \end{smallmatrix}\right] k^2} \\ x_1 \\ \vdots \\ y_5 \end{bmatrix} \tag{9.8}$$

Symmetry (9.8) is visualized in Figure 9.2.

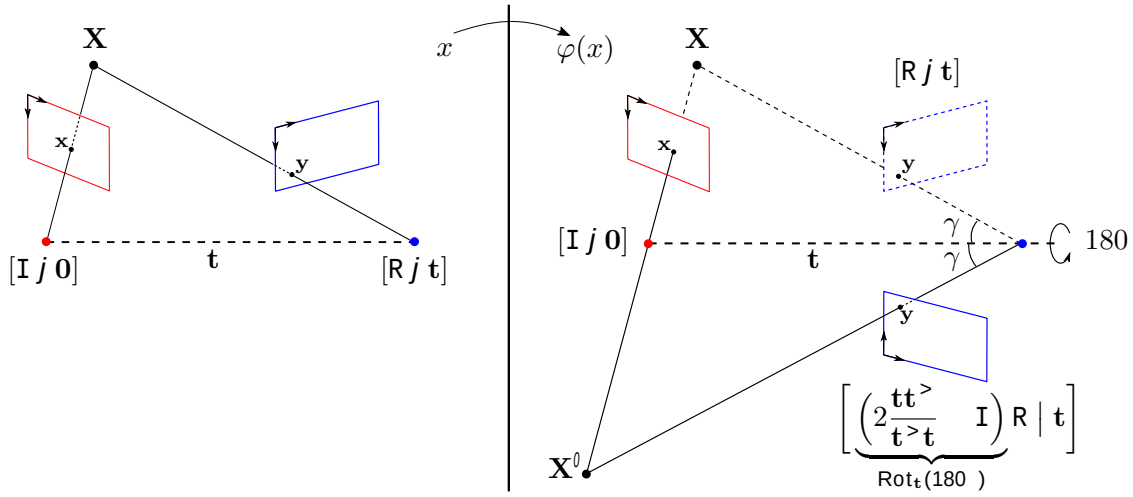


Figure 9.2

In order to find a factorizing map f_1 we can use an argument given in [3, p. 4]. We proceed as in Example 8.11. By Proposition 4.46, $\mathtt{Bir}(f)$ acts on $C(X)$ by automorphisms fixing $C(Z)$ pointwise. We can look at the fixed subfield M of this group:

$$C(Z) \qquad M = C(X)_{\mathtt{Bir}(f)} \stackrel{\text{def}}{=} \{g \in C(X) \mid \psi(g) = g \ \ \forall \psi \in \mathtt{Bir}(f)\}.$$

Since the order of $\mathtt{Bir}(f)$ equals 2, then [5, Chapter 7, Theorem 7.5.3] shows that

$$[\mathsf{C}(X) : M] = j\mathtt{Bir}(f)j = 2.$$

In other words, the field extension M $\mathsf{C}(X)$ is finite of degree 2. Thus, $M = \mathsf{C}(Y)$ for $\mathsf{C}(Y)$ from (9.3). Our task is to find a quasiprojective variety Y with a function field isomorphic to M. Since $\mathsf{C}(X)$ is finitely generated over C, then so is M. Using Reynolds operator [6, Chapter 7, §3, Definition 2], the generators $g_1, \ldots, g_r$ of M over C can be computed (for the case when $\mathtt{Bir}(f)$ is a matrix group, see [6, Chapter 7]). Then f_1 is given by $g_1, \ldots, g_r$ 2 M $\mathsf{C}(X)$ and, as was explained above, the equations defining Y can be computed using rational implicitization ([6, Chapter 3, §3]), which can be computationally hard if $g_1, \ldots, g_r$ have a complicated form.

9.2 Point-line minimal problems 3100_0 and 3010_0

Similarly, the branched covers for the problems 3100_0 (see Figure 9.3) and 3010_0 (see Figure 9.4) can be defined [11, Definition 1]. In this definition the symbols p, l, I, m for these two problems are the following: for 3100_0: $p = 4, l = 1, I = f(1,1),(2,1),(3,1)g, m = 3$ and for 3010_0: $p = 3, l = 1, I = ?, m = 3$. The branched cover $\Phi_{p,l,I,m}$ from [11, Definition 1] is denoted here by f.

9.2.1 PLMP 3100_0

The Galois group $\mathtt{Gal}(f)$ has been computed by [9] using [10]:

$$\mathtt{Gal}(f) = \Big\langle (1\ 37\ 51\ 63\ 22\ 44)(2\ 31\ 13\ 36\ 16\ 47\ 50\ 24\ 29\ 25\ 34\ 28\ 53\ 11\ 27\ 9\ 4\ 49\ 30\ 58\ 32\ 56)$$

$$(3\ 60\ 15\ 40\ 39\ 61\ 42\ 12\ 35\ 54\ 45\ 17\ 46\ 41\ 23\ 55\ 57\ 64\ 19\ 14\ 7\ 52)$$

$$(5\ 6\ 43\ 18)(8\ 21\ 33\ 48)(10\ 38\ 62\ 26\ 59\ 20),$$

$$(1\ 42\ 59\ 64)(2\ 31\ 57\ 60\ 26\ 15\ 47\ 37\ 45\ 17\ 50\ 11\ 10\ 27\ 55\ 22)$$

$$(3\ 49\ 13\ 51\ 36\ 39\ 62\ 28\ 53\ 61\ 46\ 44\ 41\ 4\ 20\ 52)$$

$$(5\ 14\ 23\ 29\ 34)(6\ 18)(7\ 21\ 25\ 16\ 19)(8\ 33)(9\ 12\ 54\ 43\ 32)(24\ 63\ 30\ 38)(35\ 40\ 58\ 56\ 48)\Big\rangle.$$

Using GAP we can compute a decomposition of the set $f1, \ldots, 64g$ into blocks of $\mathtt{Gal}(f)$ (see Definition 8.1 and Proposition 8.4). The GAP code which does that is

```
Blocks(Galois_Group,[1..64]);
```

The result of the above command is the decomposition of $f1, \ldots, 64g$ into 16 blocks of $\mathtt{Gal}(f)$ of size 4:

$$B = \Big\{ f1, 38, 59, 63g, f2, 28, 45, 52g, f3, 17, 31, 53g, f4, 39, 47, 55g, f5, 21, 43, 48g, f6, 8, 18, 33g,$$

$$f7, 34, 54, 56g, f9, 16, 23, 40g, f10, 26, 44, 51g, f11, 13, 46, 60g, f12, 19, 29, 58g, f14, 25, 32, 35g,$$

$$\{f15, 27, 36, 41g, f20, 22, 37, 62g, f24, 30, 42, 64g, f49, 50, 57, 61g\}.$$

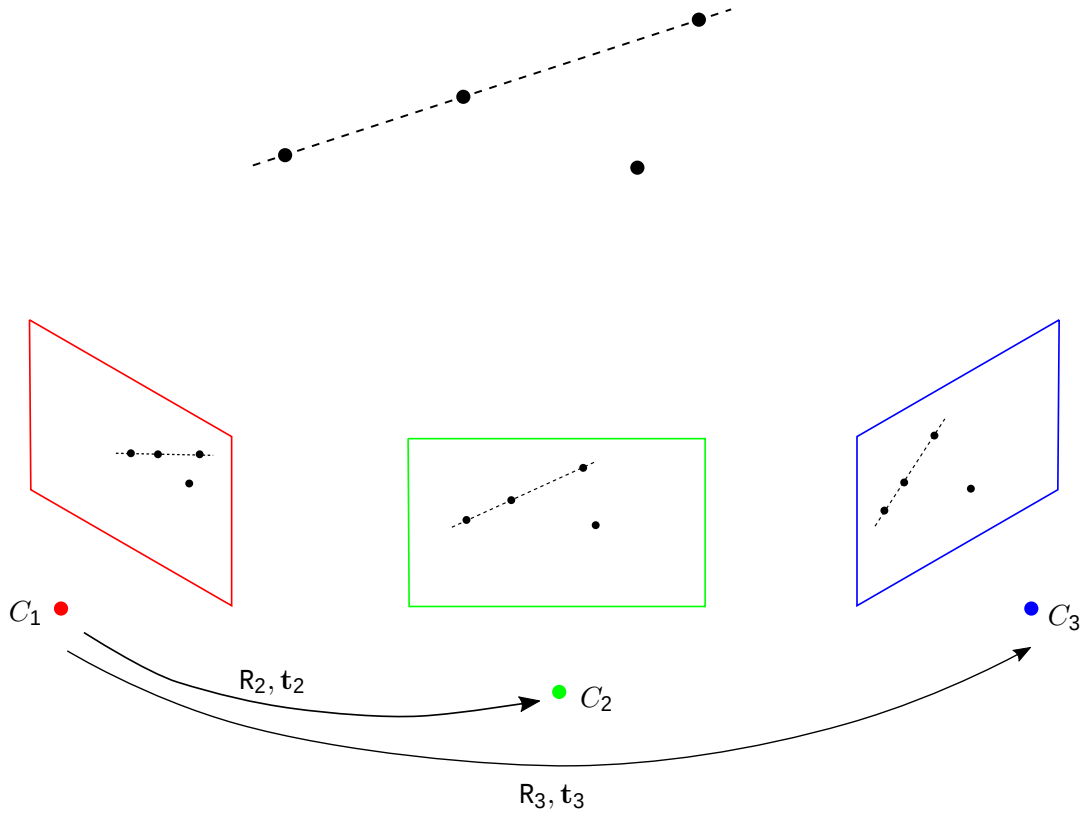


Figure 9.3

By Definition 8.2, $\mathtt{Gal}(f)$ is imprimitive. By Proposition 8.10, f is decomposable, which means that there is a quasiprojective variety Y such that f factors

$$X \xrightarrow{f_1} Y \xrightarrow{f_2} Z$$

where f_1 and f_2 are finite rational maps with degrees $d_1 = 4$ and $d_2 = \frac{64}{4} = 16$, respectively, such that $f = f_2 \quad f_1$. We check if f has symmetries (see Definition 7.19). From (7.3) we know that $\mathtt{Bir}(f)$ is isomorphic to the centralizer of $\mathtt{Gal}(f)$ in S_{64}. To find out if $\mathtt{Bir}(f)$ is nontrivial we compute the centralizer of the Galois group $\mathtt{Gal}(f)$ in S_{64} using the GAP command

```
Bir := Centralizer(SymmetricGroup(64), Galois_Group);
```

We obtain two generators of $\mathtt{Bir}(f)$ of order 2. Thus,

$$\mathtt{Bir}(f) = \mathrm{C}_{S_{64}}(\mathtt{Gal}(f)) = \mathrm{Z}/2\mathrm{Z} \quad \mathrm{Z}/2\mathrm{Z}.$$

These birational maps represent the symmetries of the problem 3100_0. We can use them to compute f_1 (as was explained in Section 9.1). Possibly, it will be computationally hard to get f_1 from $\texttt{Bir}(f)$. We haven't yet computed the factorization map f_1 for this problem, however, we believe that it can be found ad hoc. This remains an open problem for the future work.

9.2.2 PLMP 3010_0

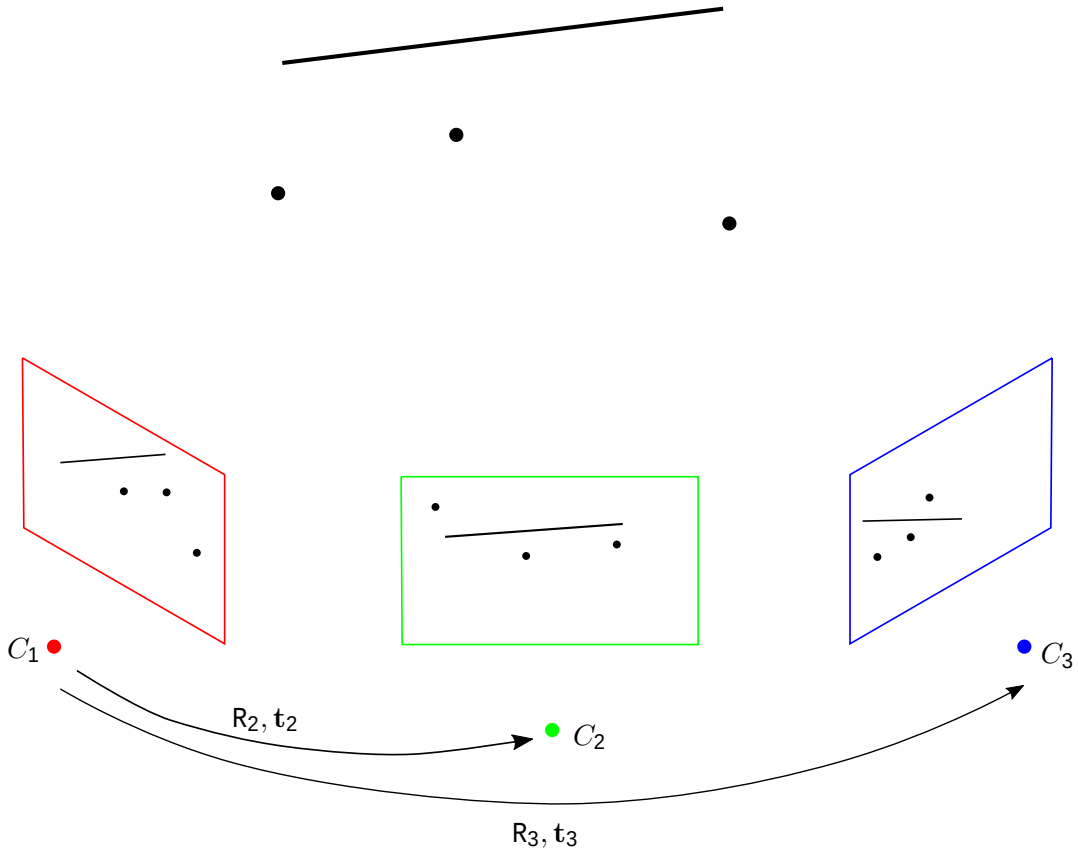


Figure 9.4

The Galois group $\texttt{Gal}(f)$ has been computed by [23] using [16]:

$$\texttt{Gal}(f) = S_{216}.$$

Thus, $\texttt{Gal}(f)$ is primitive which means, by Proposition 8.10, there doesn't exist a quasiprojective variety Y such that f factors

$$X \overset{f_1}{\dashrightarrow} Y \overset{f_2}{\dashrightarrow} Z$$

as a composition of finite rational maps f_1, f_2 of degrees $d_1, d_2 > 1$, respectively.

10 Conclusion

In this work, we have shown how the symmetries of parametric polynomial systems can be found. Such systems arise in, for example, computer vision (minimal problems) or robotics (inverse kinematics). We saw that the symmetries are encoded in the Galois/monodromy group of this polynomial system. The crucial fact is that the Galois/monodromy group can be computed using numerical algebraic geometry. Thus, there is a numerical method for revealing the symmetries.

Concerning the applications in computer vision (see Chapter 9), we have reviewed the theoretical insights of why the 5-point problem with 20 solutions is reducible to a simpler problem with 10 solutions (the essential matrix formulation). We have also shown that the point-line minimal problem 3100_0 with 64 solutions can be reduced to a simpler problem with 16 solutions. However, we haven't yet computed the equations which define the reduced problem with 16 solutions. This will be a part of the future work. We have also shown that the point-line minimal problem 3010_0 with 216 solutions cannot be reduced to a simpler problem with less number of solutions.

Bibliography

[1] C. Améndola and J. I. Rodriguez. *Solving Parameterized Polynomial Systems with Decomposable Projections*. 2016. arXiv: 1612.08807 [math.AG].

[2] C. Awtrey, N. Mistry, and N. Soltz. "Centralizers of Transitive Permutation Groups and Applications to Galois Theory". In: *Missouri J. Math. Sci.* 27.1 (Nov. 2015), pp. 16–32. doi: 10.35834/mjms/1449161364.

[3] T. Brysiewicz, J. I. Rodriguez, F. Sottile, and T. Yahl. *Solving Decomposable Sparse Systems*. 2020. arXiv: 2001.04228 [math.AG].

[4] B. Conrad. *Transcendence bases (Lecture notes)*. url: http://virtualmath1.stanford.edu/~conrad/121Page/handouts/trdeg.pdf.

[5] D. A. Cox. *Galois Theory*. Wiley-Interscience, 2004.

[6] D. A. Cox, D. O'Shea, and J. Little. *Ideals, Varieties, and Algorithms*. Springer, 2015.

[7] F. Cukierman. "Monodromy of projections". In: *Mat. Contemp.* 16 (1999), pp. 9–30.

[8] E. Dobson. *Imprimitive Permutation Groups*. 2013.

[9] T. Du□. *Personal communication*.

[10] T. Du□, C. Hill, A. Jensen, K. Lee, A. Leykin, and J. Sommars. *Solving polynomial systems via homotopy continuation and monodromy*. 2016. arXiv: 1609.08722 [math.AG].

[11] T. Du□, K. Kohn, A. Leykin, and T. Pajdla. *PLMP – Point-Line Minimal Problems in Complete Multi-View Visibility*. 2019. arXiv: 1903.10008 [cs.CV].

[12] *GAP - Groups, Algorithms, Programming*. https://www.gap-system.org.

[13] J. Harris. "Galois groups of enumerative problems". In: *Duke Math. J.* (1979), pp. 685–724. doi: 10.1215/S0012-7094-79-04635-0.

[14] J. Harris. *Algebraic Geometry: A First Course*. Springer, 1992.

[15] A. Hatcher. *Algebraic Topology*. Cambridge University Press, 2002.

[16] J. D. Hauenstein, J. I. Rodriguez, and F. Sottile. *Numerical computation of Galois groups*. 2016. arXiv: 1605.07806 [math.AG].

[17] V. Korotynskiy. "Using Symmetries in Solving Minimal Problems in Computer Vision". Bachelor's thesis. 2018.

[18] Z. Kukelova, P. Krsek, V. Smutny, and T. Pajdla. "Groebner Basis Solutions to Satellite Trajectory Control by Pole Placement". In: *Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference (AMOS'13)*. 2013.

[19] V. Larsson and K. Åström. “Uncovering symmetries in polynomial systems”. In: *ECCV 2016, Proceedings of the 14th European Conference on Computer Vision*. Amsterdam, The Netherlands: Springer Verlag, 2016, pp. 252–267. doi: 10.1007/978-3-319-46487-9_16.

[20] J. McKernan. *Classification of finitely generated field extension (Lecture notes)*. url: http://math.mit.edu/~mckernan/Teaching/07-08/Autumn/18.735/l_1.pdf.

[21] D. Nister, R. Hartley, and H. Stewenius. “Using Galois Theory to Prove Structure from Motion Algorithms are Optimal”. In: *2007 IEEE Conference on Computer Vision and Pattern Recognition*. 2007, pp. 1–8. doi: 10.1109/CVPR.2007.383089.

[22] B. C. Öner. “Galois Theory , Monodromy Groups and Flexes of Plane Cubic Curves”. Bachelor’s thesis. 2012.

[23] M. Regan. *Personal communication*.

[24] I. R. Shafarevich. *Basic Algebraic Geometry 1*. Springer, 2013.

[25] A. Sommese and C. Wampler. *The Numerical Solution of Systems of Polynomials Arising in Engineering and Science*. World Scientific, 2005. doi: https://doi.org/10.1142/5763.

[26] J. Šťovíček. *Algebraic Geometry (Lecture notes)*. url: https://www.karlin.mff.cuni.cz/~stovicek/dl/17-18-zs/alg-geom.pdf.

[27] H. Tari, H.-J. Su, and J. D. Hauenstein. “Classification and Complete Solution of the Kinetostatics of a Compliant Stewart–Gough Platform”. In: *Mechanism and Machine Theory*. 2012, pp. 177–186. doi: 10.1016/j.mechmachtheory.2011.10.011.

[28] K. Werndli. “Elementary GAGA”. Master’s thesis. 2011.